

(12)特許協力条約に基づいて公開された国際出願

(19)世界知的所有権機関
国際事務局(43)国際公開日
2004年8月19日 (19.08.2004)

PCT

(10)国際公開番号
WO 2004/071130 A1(51)国際特許分類:
3/20, H03H 21/00, 17/02, H04M 1/60

(72)発明者: および

(75)発明者/出願人(米国についてのみ): 小林 和則
(KOBAYASHI, Kazunori) [JP/JP]; 〒1808585 東京都
武蔵野市緑町三丁目9番11号 NTT 知的財産
センタ内 Tokyo (JP). 古家 賢一 (FURUYA, Kenichi)
[JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目9番
11号 NTT 知的財産センタ内 Tokyo (JP).

(21)国際出願番号: PCT/JP2004/001261

(74)代理人: 草野 卓, 外 (KUSANO, Takashi et al.); 〒
1600022 東京都新宿区新宿四丁目2番21号相模ビル
Tokyo (JP).

(22)国際出願日: 2004年2月6日 (06.02.2004)

(25)国際出願の言語: 日本語

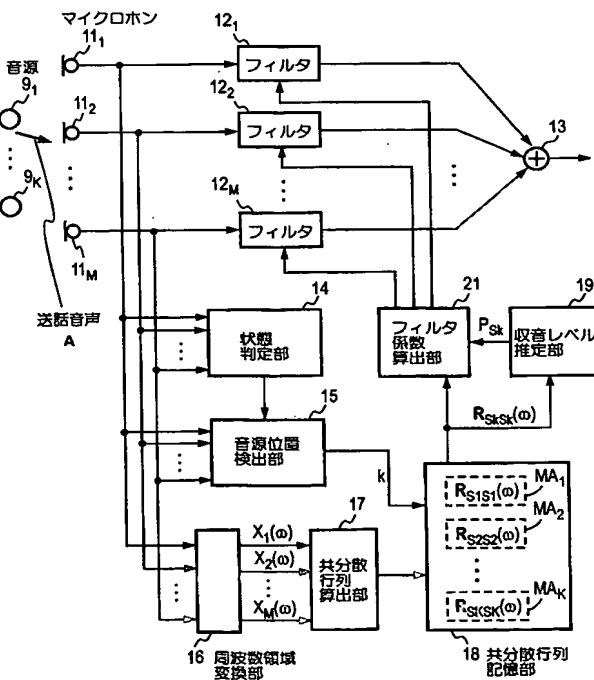
(26)国際公開の言語: 日本語

(30)優先権データ:
特願2003-030676 2003年2月7日 (07.02.2003) JP
特願2003-058626 2003年3月5日 (05.03.2003) JP(71)出願人(米国を除く全ての指定国について): 日本電
信電話株式会社 (NIPPON TELEGRAPH AND TELE-
PHONE CORPORATION) [JP/JP]; 〒1008116 東京都
千代田区大手町二丁目3番1号 Tokyo (JP).(81)指定国(表示のない限り、全ての種類の国内保護が
可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR,
BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU,
ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS,

[続葉有]

(54)Title: SOUND COLLECTING METHOD AND SOUND COLLECTING DEVICE

(54)発明の名称: 収音方法及び収音装置



9₁...SOUND SOURCE
11₁...MICROPHONE
12₁...FILTER
12₂...FILTER
12_M...FILTER

A...SPEECH TRANSMISSION SOUND
14...STATE JUDGING SECTION
15...SOUND SOURCE POSITION DETECTING SECTION
21...FILTER FACTOR CALCULATING SECTION
19...COLLECTED SOUND LEVEL DEDUCING SECTION
16...FREQUENCY RANGE CHANGING SECTION
17...COVARIANCE MATRIX CALCULATING SECTION
18...COVARIANCE MATRIX STORAGE SECTION

(57) Abstract: When a speech section is detected by a state judging section (14), the positions of the sound sources(9_1 to 9_K) are determined by a sound source position detecting section (15). A covariance matrix of the received sound signal is determined correspondingly to each sound source by a covariance matrix calculating section (17) and stored in a covariance matrix storage section (18) correspondingly to the sound source. The levels of the collected sounds from the sound sources are deduced from the covariance matrices by a collected sound level deducing section (19). The filter factors are determined from the deduced collected sound levels and the covariance matrices by a filter factor calculating section (21) so that the output level may be a predetermined value. The determined filter factors are assigned to filters (12_1 to 12_M). The received sound signals from the microphones are processed by means of the filters, and the results of the processings are added by an adder (13) and outputted as a speech transmission signal. Thus, irrespective of the positions of the sound sources, a speech transmission signal of desired level can be generated.

(57)要約: 状態判定部14により発話区間と判定すると、音源位置検出部15により各音源 9_1 ~ 9_K の位置を検出し、それらの音源に対応して共分散行列算出部17により受音信号の共分散行列を求め、各音源に対応して共分散行列記憶部18に保存する。共分散行列から各音源の収音レベルを収音レベル推定部19により推定し、出力レベルが所定値となるよう、フィルタ係数算出部21により推定収音レベルと共に分散行列からフィルタ係数を決定し、フィルタ 12_1 ~ 12_M に設定する。これらフィルタによりそれぞれのマイクロфонからの受音信号を処理し、処理結果を加算器13で加算して送話信号として出力することにより、音源の位置にかかわらず、それぞれ所望のレベルの送話信号を生成できる。

WO 2004/071130 A1



LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA,
NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE,
SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, YU, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL,
SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG,
KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH,
CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU,

MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG,
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 国際調査報告書

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

明細書

収音方法及び収音装置

技術分野

この発明は収音方法及び収音装置に関し、特に、複数の音源からの送話音声を収音し、それぞれの音量を調整して出力する収音方法及び収音装置に関する。

従来の技術

例えば互いに離れた複数地点間で行う遠隔会議(teleconference)においては、各地点で異なる位置の複数の参加者の音声を1つのマイクロホンで収音すると、参加者ごとにマイクロホンまでの距離、及び発声音量が異なるため、受音信号レベルが大きく異なってしまう。遠隔地点の受信側では、送信側の参加者によって再生音声の音量が大きく異なり、非常に聞き取りにくい場合がある。

図17は例えば日本国特許出願公開 8-250944 に示されている従来の収音装置の原理的構成をプロック図で示す。この従来の収音装置は、マイクロホン41と、パワー算出部42と、増幅率設定部43と、増幅器44により構成される。パワー算出部42は、マイクロホン41で受音された信号の長時間平均パワー P_{ave} を求める。長時間平均パワーは、受音信号を二乗し、時間積分することで求めることができる。次に、増幅率設定部43は、パワー算出部42で算出された受音信号の長時間平均パワー P_{ave} と、あらかじめ設定した所望の送出レベル P_{opt} に基づき、増幅率Gを設定する。増幅率Gは、例えば、次式(1)により求められる。

$$G = (P_{opt}/P_{ave})^{1/2} \quad (1)$$

増幅器44は、設定された増幅率Gで、マイクロホン受音信号を増幅し、出力する。

以上の処理により、出力信号パワーは、所望の送出レベル P_{opt} となり、音量が自動的に調整される。しかし、従来技術の収音方法では、長時間平均パワーを基に増幅率を決定しているので、適切な増幅率が設定されるまでに数秒から数十秒の遅延が生じる。これにより、話者が複数人存在し、それぞれの話者音声が異なるレベルでマイクロホンに収音される場合、発話者が切り替るたびに、増幅率設定の適応が遅れ、不適切な音量となるという問題が生じる。

この発明の目的は、話者が複数人存在し、それぞれの話者音声が異なるレベルで

マイクロホンに収音される場合においても、各話者音声が適切な音量となるように音量を自動調整する収音装置、収音方法、及びそのプログラムを提供することである。

5 発明の開示

この発明による複数チャネルのマイクロホンにより音源から発せられた音を収音する収音方法は、

(a) 上記複数チャネルのマイクロホンで受音された受音信号から、発話区間を判定する発話判定段階を含む状態判定段階と、

10 (b) 上記発話判定段階において発話区間と判定された場合に、上記受音信号から、各音源位置を検出する音源位置検出段階と、

(c) 上記受音信号を周波数領域の信号に変換する周波数領域変換段階と、

(d) 上記周波数領域の受音信号の共分散行列を算出する共分散行列算出段階と、

15 (e) 上記音源位置検出段階の検出結果に基づき、各音源に対し上記共分散行列を記憶する共分散行列記憶段階と、

(f) 記憶されている上記共分散行列と予め決めた出力レベルに基づき上記複数チャネルのフィルタ係数を算出するフィルタ係数算出段階と、

(g) 上記複数チャネルの受音信号をそれぞれ上記複数チャネルのフィルタ係数でフィルタリングするフィルタリング段階と、

20 (h) 上記複数チャネルのフィルタリング結果を加算して送話信号として出力する加算段階、
とを含む収音方法。

この発明によれば、音響空間に配置された複数チャネルのマイクロホンにより各音源から発せられた音を収音する収音装置は、

複数チャネルのマイクロホンの各々で受音された受音信号から、発話区間を判定する発話判定部を含む状態判定部と、

上記発話判定部において発話区間と判定された場合に、上記受音信号から、音源位置を検出する音源位置検出部と、

上記受音信号を周波数領域の信号に変換する周波数領域変換部と、

上記周波数領域の複数チャネルの受音信号の共分散行列を各音源ごとに算出する
共分散行列算出部と、

上記音源位置検出手段の検出結果に基づき、上記共分散行列を音源ごとに記憶す
5 る共分散行列記憶部と、

記憶された上記共分散行列を用いて、各音源に対する送話信号レベルが各々所望
のレベルとなるよう複数チャネルのフィルタ係数を算出するフィルタ係数算出部と、

上記マイクロホン受音信号を、上記複数チャネルのフィルタ係数で、各々フィル
タリングする複数チャネルのフィルタと、

10 上記複数チャネルのフィルタの出力信号を加算し、送話信号として出力する加算
器、

とを含むように構成される。

この発明の第2の観点によれば、受話信号がスピーカから再生される音響空間内
で少なくとも1つのマイクロホンにより少なくとも1つの音源からの発音を收音す
15 る収音方法は、

(a) 上記少なくとも1チャネルのマイクロホンで受音された受音信号と受話信号か
ら発話区間及び受話区間を判定する状態判定段階と、

(b) 上記受音信号及び上記受話信号を周波数領域の信号に変換する周波数領域変
換段階と、

20 (c) 上記周波数領域の受音信号及び受話信号から発話区間での共分散行列及び受
話区間での共分散行列を算出する共分散行列算出段階と、

(d) 上記状態判定段階の判定結果に基づいて、上記共分散行列を上記発話区間及
び上記受話区間ごとに記憶する共分散行列記憶段階と、

(e) 上記発話区間及び上記受話区間での記憶された共分散行列に基づいて受音信号

25 に含まれる受話信号成分である音響エコーを抑圧するよう上記少なくとも1チャネ
ルの受音信号用のフィルタ係数と上記受話信号用のフィルタ係数を算出するフィル
タ係数算出段階と、

(f) 上記受話信号及び上記受音信号を、上記受話信号用のフィルタ係数と上記少
くとも1チャネルの受音信号用のフィルタ係数でフィルタリングするフィルタリン

グ段階と、

(g) 上記フィルタリングされた信号を加算して送話信号として出力する加算段階、
とを含む。

この発明の第2の観点による収音装置は、

5 音源からの送話音を収音して受音信号を出力する少なくとも1チャネルのマイク
ロホンと、

受話信号を再生出力するスピーカと、

受音信号と受話信号から発話区間及び受話区間を判定する状態判定部と、

上記受話信号及び上記受音信号を周波数領域の信号に変換する周波数領域変換部

10 と、

上記発話区間及び受話区間ごとに上記周波数領域の受音信号と受話信号の共分散
行列を算出する共分散行列算出部と、

上記共分散行列を発話区間及び受話区間ごとに記憶する共分散行列記憶部と、

記憶された上記共分散行列に基づいて上記受話信号の音響エコーを抑圧するよう

15 に上記少なくとも1チャネルの受音信号用のフィルタ係数と上記受話信号用のフィ
ルタ係数を算出するフィルタ係数算出部と、

上記受音信号用のフィルタ係数と上記受話信号用のフィルタ係数が設定され、上
記受音信号と上記受話信号をフィルタリングして出力する受音信号用フィルタと受
話信号用フィルタと、

20 上記受音信号用フィルタの出力と上記受話信号用フィルタの出力を加算し、加算
結果を送話信号として出力する加算器、
とを含むように構成される。

本発明は、話者が複数人存在し、それぞれの話者音声が異なるレベルでマイクロ
ホンに収音される場合においても、複数のマイクロホンを用いて指向特性を適切に
25 制御することにより、話者ごとに適切な音量となるように音量を自動調整する。

図面の簡単な説明

図1は本発明の第1の実施例である収音装置を示すブロック図。

図2は図1における状態判定部1~4の構成例を示すブロック図。

図3は図1における音源位置検出部15の構成例を示すブロック図。

図4は図1におけるフィルタ係数算出部21の構成例を示すブロック図。

図5は図1の収音装置を使用した収音方法の第1の例を示すフロー図。

図6は図1の収音装置を使用した収音方法の第2の例を示すフロー図。

5 図7は図1の収音装置を使用した収音方法の第3の例を示すフロー図。

図8は本発明の第2の実施例である収音装置を示すブロック図。

図9は図8における状態判定部14の構成例を示すブロック図。

図10は本発明の第3の実施例である収音装置を示すブロック図。

10 図11は図7における状態判定部14の構成例を示すブロック図。

図12は本発明の第4実施例の収音装置を示すブロック図。

図13は本発明の第5実施例の収音装置を示すブロック図。

図14は図4における重み係数設定部21Hの構成例を示すブロック図。

図15は図4における重み係数設定部21Hの他の構成例を示すブロック図。

図16は図4に設ける白色化部21Jの構成例を示すブロック図。

15 図17は各実施例において、共分散行列の平均化機能を設けた場合の共分散行列記憶部18の構成例を示すブロック図。

図18Aは第1実施例による処理を行う前の話者A及びBのシミュレーションによる音声波形を示す図。

図18Bは第1実施例による処理を行った場合の話者A及びBのシミュレーションによる音声波形を示す図。

20 図19は第3実施例による音響エコー及び雑音の抑圧を示すシミュレーションによる受話及び送話音声波形を示す図。

図20は従来の収音装置を示すブロック図。

25 発明を実施するための最良の形態

第1実施例

図1は、本発明の第1の実施例である収音装置のブロック図である。

本実施例の収音装置は、音響空間内に配置されたMチャネルのマイクロホン11₁～11_Mと、フィルタ12₁～12_Mと、加算器13と、状態判定部14と、音源

位置検出部15と、周波数領域変換部16と、共分散行列算出部17と、共分散行列記憶部18と、収音レベル推定部19と、フィルタ係数算出部21により構成される。

本実施例は、音響空間内の音源 $9_1 \sim 9_K$ の位置を検出し、各音源位置に対する受音信号の周波数領域での共分散行列を求めて保存しておき、これらを用いてフィルタ係数を求める。これらのフィルタ係数で、マイクロホン受音信号をフィルタリングすることで、各音源からの信号が一定の音量となるように制御する。この発明の実施例において、マイクロホン $11_1 \sim 11_M$ の各出力受音信号は特に図示しないがデジタルアナログ変換器により予め決められたサンプリング周波数でデジタル値に変換して得られたデジタルの受音信号であるものとする。他の実施例においても同様である。

まず、状態判定部14では、マイクロホン $11_1 \sim 11_M$ で受音したマイクロホン受音信号から、発話区間を検出する。例えば、図2に示すように、状態判定部14ではマイクロホン $11_1 \sim 11_M$ からの全受話信号を加算部14Aで加算し、その加算出力を短時間平均パワー算出部14Bと長時間平均パワー算出部14Cでそれぞれ短時間平均パワー（例えば0.1～1s間程度の平均パワー） P_{avS} と、長時間平均パワー（例えば1～100s間程度の平均パワー） P_{avL} を求め、割算部14Dで短時間平均パワーと長時間平均パワーの比 $R_p = P_{avS}/P_{avL}$ を求め、発話判定部14Eでそのパワー比 R_p を予め決めた発話の閾値 R_{thU} と比較し、それを超える場合に発話区間と判定する。

状態判定部14の判定結果が発話区間であった場合に、音源位置検出部15は、音源の位置を推定する。音源位置の推定方法は、例えば相互相関法による方法がある。

マイクロホンの数をM（Mは2以上の整数）個とし、i番目マイクロホン 11_i とj番目マイクロホン 11_j で受音された信号より求められる受音信号間遅延時間差の測定値を τ_{ij} とする。受音信号間遅延時間差の測定値は、信号間の相互相関を求め、その最大ピーク位置から求めることができる。次に、m(m=1,...,M)番目のマイクロホン 11_m の受音位置を (x_m, y_m, z_m) 、推定音源位置を $(\hat{x}, \hat{y}, \hat{z})$ と表す。これらの位置から求められる受音信号間遅延時間差の測定値 $\hat{\tau}_{ij}$ は、式(2)で表される。

$$\hat{\tau}_{ij} = \frac{1}{c} \sqrt{(x_i - \hat{X})^2 + (y_i - \hat{Y})^2 + (z_i - \hat{Z})^2} - \frac{1}{c} \sqrt{(x_j - \hat{X})^2 + (y_j - \hat{Y})^2 + (z_j - \hat{Z})^2} \quad (2)$$

ただし、 c は音速である。

次に、受音信号間遅延時間差の測定値 τ_{ij} 及び推定値 $\hat{\tau}_{ij}$ に音速 c を乗じ距離に換算したものを、それぞれ受音位置間距離差の測定値 d_{ij} 及び推定値 \hat{d}_{ij} とし、これら 5 の二乗平均誤差 $e(\mathbf{q})$ を求めれば、式 (3) となる。

$$\begin{aligned} e(\mathbf{q}) &= \sum_{i=1}^{M-1} \sum_{j=i+1}^M |d_{ij} - \hat{d}_{ij}|^2 \\ &= \sum_{i=1}^{M-1} \sum_{j=i+1}^M \left| d_{ij} - \sqrt{(x_i - \hat{X})^2 + (y_i - \hat{Y})^2 + (z_i - \hat{Z})^2} + \sqrt{(x_j - \hat{X})^2 + (y_j - \hat{Y})^2 + (z_j - \hat{Z})^2} \right|^2 \\ &= \sum_{i=1}^{M-1} \sum_{j=i+1}^M |d_{ij} - r_i + r_j|^2 \end{aligned} \quad (3)$$

ただし、 $\mathbf{q} = (\hat{X}, \hat{Y}, \hat{Z})$ である。 r_i 及び r_j は推定音源位置 $\mathbf{q} = (\hat{X}, \hat{Y}, \hat{Z})$ とマイクロホン 10 11_i、11_j 間の距離を表している。

式 (3) の二乗平均誤差 $e(\mathbf{q})$ を最小化する解を求めれば、受音信号間遅延時間差の測定値と推定値の誤差が最小となる推定音源位置を求めることができる。ただし、式 (3) は非線形連立方程式となっており、解析的に解くことは困難であるので、逐次修正を用いた数値解析により求める。

式 (3) を最小化する推定音源位置 $(\hat{X}, \hat{Y}, \hat{Z})$ を求めるには、式 (3) のある点における勾配を求め、誤差が小さくなる方向に、推定音源位置を修正していく、勾配が 0 となる点を求めればよいので、推定音源位置の修正は次式 (4) を $u=0, 1, \dots$ について繰り返し演算する。

$$\mathbf{q}_{(u+1)} = \mathbf{q}_{(u)} - \alpha \cdot \text{grad } e(\mathbf{q}) \Big|_{\mathbf{q}=\mathbf{q}_{(u)}} \quad (4)$$

ただし、 α は修正のステップサイズであり、 $\alpha > 0$ の値を設定する。 $\mathbf{q}_{(u)}$ は、 u 回修正後の \mathbf{q} を表し、 $u=0$ の場合の $\mathbf{q}_{(0)} = (\hat{X}_0, \hat{Y}_0, \hat{Z}_0)$ を予め決めた任意の初期値とする。grad は勾配を表し、次式 (5) ~ (10) で表される。

$$\text{grad } e(\mathbf{q}) = \left(\frac{\partial e(\mathbf{q})}{\partial \hat{X}}, \frac{\partial e(\mathbf{q})}{\partial \hat{Y}}, \frac{\partial e(\mathbf{q})}{\partial \hat{Z}} \right) \quad (5)$$

$$\frac{\partial e(q)}{\partial \hat{X}} = 2 \sum_{i=1}^{M-1} \sum_{j=i+1}^M \{d_{ij} - r_i + r_j\} \cdot \left\{ \frac{x_i - \hat{X}}{r_i} - \frac{x_j - \hat{X}}{r_j} \right\} \quad (6)$$

$$\frac{\partial e(q)}{\partial \hat{Y}} = 2 \sum_{i=1}^{M-1} \sum_{j=i+1}^M \{d_{ij} - r_i + r_j\} \cdot \left\{ \frac{y_i - \hat{Y}}{r_i} - \frac{y_j - \hat{Y}}{r_j} \right\} \quad (7)$$

$$\frac{\partial e(q)}{\partial \hat{Z}} = 2 \sum_{i=1}^{M-1} \sum_{j=i+1}^M \{d_{ij} - r_i + r_j\} \cdot \left\{ \frac{z_i - \hat{Z}}{r_i} - \frac{z_j - \hat{Z}}{r_j} \right\} \quad (8)$$

$$r_i = \sqrt{(x_i - \hat{X})^2 + (y_i - \hat{Y})^2 + (z_i - \hat{Z})^2} \quad (9)$$

5 $r_j = \sqrt{(x_j - \hat{X})^2 + (y_j - \hat{Y})^2 + (z_j - \hat{Z})^2} \quad (10)$

以上、式(4)を繰返し計算することで、誤差が最小となる推定音源位置 $q = (\hat{X}, \hat{Y}, \hat{Z})$ を求めることができる。

図3は音源位置検出部15の機能構成をブロック図で示す。この例では、音源位置検出部15は、遅延時間差測定部15Aと、乗算器15Bと、距離計算部15Cと、二乗平均誤差計算部15Dと、勾配計算部15Eと、比較判定部15Fと、推定位置更新部15Gとから構成されている。

遅延時間差測定部15Aは1つの音源9_kから発音されている期間に

$$i=1, 2, \dots, M-1;$$

$$j=i+1, i+2, \dots, M$$

15 の全ての組の(i,j)についてマイクロホン11_i、11_jによる受音信号に基いて相互相関法により遅延時間差 τ_{ij} を測定する。乗算器15Bは、測定された各遅延時間差 τ_{ij} に音速cを乗算して音源とマイクロホン11_i、11_j間の距離差 d_{ij} を求める。距離計算部15Cは推定位置更新部15Gから与えられる推定音源位置($\hat{X}, \hat{Y}, \hat{Z}$)とマイクロホン11_i、11_j間の距離 r_i, r_j を式(9)、(10)により計算する。ただし、推定位置更新部15Gは、最初の推定音源位置として任意の初期値($\hat{X}_0, \hat{Y}_0, \hat{Z}_0$)を距離計算部15Cに与える。二乗平均誤差計算部15Dは前述の全ての組の(i,j)について d_{ij}, r_i, r_j を使って式(3)により二乗平均誤差 $e(q)$ を計算する。勾配計算部15Fで現在の推定音源位置と d_{ij}, r_i, r_j を使って式(6)、(7)、(8)により二乗平均誤差 $e(q)$ の勾配 $\text{grad } e(q)$ を計算する。

25 比較判定部15Fは二乗平均誤差の勾配 $\text{grad } e(q)$ の各要素と予め決めた閾値 e_{th} と

を比較し、全ての要素が閾値 e_{th} より小さくなつたか判定し、小さければそのときの推定音源位置 $(\hat{X}, \hat{Y}, \hat{Z})$ を出力する。 e_{th} より小さくなければ、推定位置更新部 15 G は勾配 $\text{grad } e(q)$ と現在の推定位置 $q = (\hat{X}, \hat{Y}, \hat{Z})$ を使って式(4)により推定位置を更新し、更新された推定位置 $q_{u+1} = (\hat{X}, \hat{Y}, \hat{Z})$ を距離計算部 15 C に与える。距離計算部 5 15 C は更新された推定位置 $(\hat{X}, \hat{Y}, \hat{Z})$ と d_{ij} を使って前述と同様に更新された r_i, r_j を計算し、以下、二乗誤差平均部 15 D は $e(q)$ を更新し、勾配計算部 15 E は更新された勾配 $\text{grad } e(q)$ を計算し、比較判定部 15 F は更新された二乗平均誤差 $e(q)$ が閾値 e_{th} より小となつたか判定する。

10 このようにして推定位置 $(\hat{X}, \hat{Y}, \hat{Z})$ の更新を二乗平均誤差の勾配 $\text{grad } e(q)$ の全要素が十分小さい値 (e_{th} より小) になるまで繰り返し実行して音源 9_k の位置 $(\hat{X}, \hat{Y}, \hat{Z})$ を推定する。他の音源についても同様に位置を推定する。

15 周波数領域変換部 16 は、各マイクロホンの受音信号を周波数領域信号に変換する。例えば受音信号のサンプリング周波数は 16kHz であり、各マイクロホン 11 $m (m=1, \dots, M)$ からの受音信号サンプルを 1 フレーム 256 サンプルごとに FFT (フアーストフーリエ変換) 処理することにより同数の周波数領域信号サンプル $X_m(\omega)$ が得られる。

20 次に、共分散行列算出部 17 では、マイクロホン受音信号の共分散を求め、共分散行列を生成する。各音源 9_k に対し周波数領域変換部 16 で得られたマイクロホン受音信号の周波数領域変換信号を $X_1(\omega) \sim X_M(\omega)$ とすると、これらの信号の $M \times M$ 共分散行列 $R_{XX}(\omega)$ は一般に式(11)で表される。

$$\begin{aligned} R_{XX}(\omega) &= \begin{pmatrix} X_1(\omega) \\ \vdots \\ X_M(\omega) \end{pmatrix} \begin{pmatrix} X_1(\omega)^* & \cdots & X_M(\omega)^* \end{pmatrix} \\ &= \begin{pmatrix} X_1(\omega)X_1(\omega)^* & X_1(\omega)X_2(\omega)^* & \cdots & X_1(\omega)X_M(\omega)^* \\ X_2(\omega)X_1(\omega)^* & X_2(\omega)X_2(\omega)^* & \cdots & X_2(\omega)X_M(\omega)^* \\ \vdots & \vdots & \ddots & \vdots \\ X_M(\omega)X_1(\omega)^* & X_M(\omega)X_2(\omega)^* & \cdots & X_M(\omega)X_M(\omega)^* \end{pmatrix} \quad (11) \end{aligned}$$

ただし、*は複素共役を表す。

次に、共分散行列記憶部 18 では、音源位置検出部 15 の検出結果に基づき、共

分散行列 $\mathbf{R}_{XX}(\omega)$ を、各音源 9_k に対する受音信号の $M \times M$ 共分散行列 $\mathbf{R}_{SkSk}(\omega)$ として保存する。

各音源 9_k ごとの M チャネルの受音信号に対する重み付けミキシングベクトルを $\mathbf{A}_k(\omega) = (a_{k1}(\omega), \dots, a_{kM}(\omega))$ で表すと、収音レベル推定部 1 9 は、各音源の収音レベル P_{Sk} を、共分散行列記憶部 1 8 に記憶した各音源 9_k に対する受音信号の共分散行列 $\mathbf{R}_{SkSk}(\omega)$ を使って次式 (12)

$$P_{Sk} = \frac{1}{W} \sum_{\omega=0}^W \mathbf{A}_k(\omega)^H \mathbf{R}_{SkSk}(\omega) \mathbf{A}_k(\omega) \quad (12)$$

により求める。ここでは重み付けミキシングベクトルは周波数特性も制御可能なベクトル $\mathbf{A}_k(\omega) = (a_{k1}(\omega), \dots, a_{kM}(\omega))$ として表しているが、周波数特性の制御を行わないのであれば、ベクトル \mathbf{A}_k の要素は予め決めた値 $a_{k1}, a_{k2}, \dots, a_{kM}$ としてよい。例えば各音源 9_k に対する重み付けミキシングベクトル \mathbf{A}_k の要素はその音源 9_k に近いマイクロホンに対応する要素ほど大きな値を与える。極端な例では、 $\mathbf{A}_k = (0, \dots, 0, a_{km}=1, 0, \dots, 0)$ のように、音源 9_k に最も近いマイクロホン 1 1 m に対応する要素のみを 1 とし、その他を全て 0 とすることもできる。以下の説明においては、式を簡略化するため、 $a_{k1}(\omega), \dots, a_{kM}(\omega)$ を単に a_{k1}, \dots, a_{kM} と表すこととする。

式 (12) における H は複素共役転置を表し、 $\mathbf{A}_k(\omega)^H \mathbf{R}_{SkSk}(\omega) \mathbf{A}_k(\omega)$ は次式のように展開することができる。

$$\begin{aligned} \mathbf{A}_k(\omega)^H \mathbf{R}_{SkSk}(\omega) \mathbf{A}_k(\omega) &= a_{k1}^*(a_{k1}X_1(\omega)X_1(\omega)^* + a_{k2}X_2(\omega)X_1(\omega)^* + \dots + a_{kM}X_M(\omega)X_1(\omega)^*) \\ &\quad + a_{k2}^*(a_{k1}X_1(\omega)X_2(\omega)^* + a_{k2}X_2(\omega)X_2(\omega)^* + \dots + a_{kM}X_M(\omega)X_2(\omega)^*) \\ &\quad \vdots \quad \vdots \quad \vdots \\ &\quad + a_{kM}^*(a_{k1}X_1(\omega)X_M(\omega)^* + a_{k2}X_2(\omega)X_M(\omega)^* + \dots + a_{kM}X_M(\omega)X_M(\omega)^*) \\ &= \Omega(\omega) \end{aligned} \quad (13)$$

式 (12) はこの式 (13) で表す $\Omega(\omega)$ が表すパワースペクトルサンプル値を、周波数領域変換部 1 6 により生成される周波数領域信号の帯域 0~W (サンプル数) に渡って加算し、Wで割り算することにより受音信号の平均パワー P_{Sk} を求めることを意味している。

例えば、音源 9_1 に最も近いマイクロホンを 1 1 $_1$ とすると、マイクロホン 1 1 $_1$ (第 1 チャネル) の受音信号に最大の重み付けを与えるように重み係数 a_{k1} の値を決

め、その他のチャネルの受音信号に対する重み係数 $a_{k2}, a_{k3}, \dots, a_{kM}$ は a_{k1} より小さい値に決める。このような重み付けを行うことにより、重み付けを行わない場合より、音源 9_1 からの受音信号の S/N を高め、あるいは室内残響の影響を減らすことができる。即ち、各音源 9_k に対する重み付けミキシングベクトル $A_k(\omega)$ の重み付け係数 5 は、例えばその音源 9_k に対応する出力音声信号の S/N が大きく、かつ室内残響が小さくなるように、マイクロホンの指向性や配置と音源の配置により最適な値を実験的に予め決めておく。しかしながら、この発明によれば全てのチャネルに等しく重み付けを行った場合でも、それぞれの音源からの受音信号を所望のレベルに制御して出力することができる。

10 次に、フィルタ係数算出部 2.1 では、各音源から発せられた音を所望の音量で収音するためのフィルタ係数を計算する。まず、各マイクロホンに接続されたのフィルタ $12_1 \sim 12_M$ のフィルタ係数を周波数領域に変換したものを $H_1(\omega) \sim H_M(\omega)$ とする。次に、これらのフィルタ係数を次式 (14) により行列としたものを $H(\omega)$ とする。

$$15 \quad H(\omega) = \begin{pmatrix} H_1(\omega) \\ \vdots \\ H_M(\omega) \end{pmatrix} \quad (14)$$

また、 k 番目音源 9_k が発音している期間の各マイクロホンで受音した信号の周波数領域変換信号を $X_{Sk,1} \sim X_{Sk,M}$ とする。

ここで、フィルタ係数行列 $H(\omega)$ に要求される条件は、マイクホン受音信号をフィルタ係数行列 $H(\omega)$ でそれぞれフィルタリングし、フィルタリング後の信号を加算したときに、各音源の信号成分のレベルが所望のレベル P_{opt} となっていることである。従って、各音源 9_k の信号をフィルタリングおよび加算した信号が、マイクロホン $11_1 \sim 11_M$ の受音信号に対する重み付けミキシングベクトル $A_k(\omega)$ と所望の利得を乗じた信号となる次式 (15) が理想条件となる。

$$(X_{Sk,1}(\omega) \cdots X_{Sk,M}(\omega))H(\omega) = \sqrt{\frac{P_{opt}}{P_{Sk}}} (X_{Sk,1}(\omega) \cdots X_{Sk,M}(\omega))A_k(\omega) \quad (15)$$

25 ただし、 $k=1, \dots, K$ であり、 K は音源の数を表す。

次に、式 (15) の条件をフィルタ係数行列 $H(\omega)$ について最小二乗解で解けば、

次式(16)となる。

$$H(\omega) = \left\{ \sum_{k=1}^K C_{Sk} R_{SkSk}(\omega) \right\}^{-1} \sum_{k=1}^K C_{Sk} \sqrt{\frac{P_{opt}}{P_{Sk}}} R_{SkSk}(\omega) A_k(\omega) \quad (16)$$

ただし、 C_{Sk} は k 番目の音源位置に対する感度拘束を与える重み係数である。感度拘束とは音源位置に対する本収音装置の周波数特性をフラットにすることを意味する。この値が大きくなればその音源に対する感度拘束が強くなり、よりフラットな周波数特性での収音が可能となるが、その他の音源位置に対する周波数特性の劣化が大きくなる。そのため、通常は C_{Sk} を 0.1~1.0 程度の値とし、全ての音源をバランスよく感度拘束するのが好ましい。

図4は式(16)で表されるフィルタ係数を計算するフィルタ係数算出部21の機能構成ブロック図を示す。この例では、共分散行列記憶部18から与えられたそれぞれの音源 $9_1 \sim 9_K$ に対応する共分散行列 $R_{S1S1} \sim R_{SKSK}$ は乗算器21A1~21AK でそれぞれ重み係数設定部21Hにより設定された重み係数 $C_{S1} \sim C_{SK}$ と乗算される。また収音レベル推定部19により推定された音源 $9_1 \sim 9_K$ に対する収音レベル $P_{S1} \sim P_{SK}$ は平方比算出部21B1~21BK で予め決めた所望の出力レベル P_{opt} との平方比 $(P_{opt}/P_{S1})^{1/2} \sim (P_{opt}/P_{SK})^{1/2}$ が演算され、演算結果はそれぞれ乗算器21C1~21CK で乗算器21A1~21AK の出力と乗算される。乗算器21C1~21CK の出力は乗算器21D1~21DK で更に重み付けミキシングベクトル $A_1(\omega) \sim A_K(\omega)$ と乗算され、これらの乗算結果の総和の行列が加算器21Eで計算される。一方、乗算器21A1~21AK の出力の総和の行列が加算器21Fで計算され、逆行列乗算器21Gにより、加算器21Fの出力である総和の行列の逆行列を加算器21Eの出力と乗算することによりフィルタ係数 $H(\omega)$ が求まる。

次に、フィルタ係数算出部21により求められたフィルタ係数 $H_1(\omega), H_2(\omega), \dots, H_M(\omega)$ は、フィルタ12₁~12_Mに設定され、マイクロホン11₁~11_Mからの受音信号をそれぞれフィルタ処理する。フィルタ処理後の信号は、加算器13で加算され、出力信号として出力される。

この発明による収音装置を使用する方法の例として、以下に3つの方法を示す。

第1の方法は、図5に示すように、まず、ステップS1で音源数KをK=0に初期設定する。次にステップS2で状態判定部14により定期的に発話検出を行い、発

話が検出されるとステップ S 3 で音源位置検出部 1 5 により音源位置検出を実行する。ステップ S 4 で、検出された音源位置が以前に検出された音源位置のいずれかと一致するか判定し、一致するものがあればステップ S 5 でその音源位置に対応する共分散行列 $R_{xx}(\omega)$ を共分散行列算出部 1 7 で新たに計算し、ステップ S 6 で共分散行列記憶部 1 8 の対応する領域の共分散行列を更新する。

ステップ S 4 で以前に検出された音源位置と一致しなかった場合は、ステップ S 7 で K を 1 だけ増加させ、ステップ S 8 でその音源位置に対応する共分散行列 $R_{xx}(\omega)$ を共分散行列算出部 1 7 で新たに計算し、ステップ S 9 で共分散行列記憶部 1 8 の新たな領域に共分散行列を記憶する。

10 次に、ステップ S 1 0 で記憶されている共分散行列から、収音レベル推定部 1 9 により収音レベルを推定し、ステップ S 1 1 で推定収音レベルと共に分散行列を使ってフィルタ係数算出部 2 1 によりフィルタ係数 $H_1(\omega) \sim H_M(\omega)$ の算出を行い、ステップ S 1 2 でフィルタ $H_{1,1} \sim H_{M,M}$ の設定フィルタ係数を更新する。

第 2 の方法は、図 6 に示すように、ステップ S 1 で予め音源数の最大値を K_{max} に 15 設定し、また音源数 K の初期値を 0 に設定しておく。以下のステップ S 2 から S 6 は図 5 の場合と同様に発話検出を行い、発話が検出されると音源位置の検出を行い、検出した音源位置が過去に検出したものと一致するか判定し、一致するものがあれば共分散行列を計算して対応する領域に更新保存する。

ステップ S 4 で検出音源位置が過去の検出音源位置と一致しない場合は、ステップ S 7 で K に 1 を加算し、ステップ S 8 で K が最大値 K_{max} より大となったか判定する。最大値 K_{max} を越えてなければステップ S 9 で検出位置に対する共分散行列を計算し、ステップ S 1 0 で新たな領域に共分散行列を保存する。ステップ S 8 で K が最大値 K_{max} を超えている場合は、ステップ S 1 1 で $K = K_{max}$ とし、ステップ S 1 2 で共分散行列記憶部 1 8 に記憶されている共分散行列のうち、最も過去に更新された共分散行列を消去し、その領域に、ステップ S 1 3 で共分散行列算出部 1 7 により計算された新たな共分散行列をステップ S 1 4 で格納する。以下のステップ S 1 5, S 1 6, S 1 7 は図 5 のステップ S 1 0, S 1 1, S 1 2 と同様に、共分散行列から各音源の推定収音レベルを求め、フィルタ係数を計算してフィルタ $H_{1,1} \sim H_{M,M}$ に設定する。この方法では、音源数 K の最大値を K_{max} に制限することにより図 5 の

方法に比べて共分散行列記憶部 18 の記憶領域を少なくすることができる利点がある。

上記第 1 及び第 2 の方法では、音源の発話を検出するごとに必ず共分散行列の計算と保存、及びフィルタ係数の更新を行う場合を示したが、第 3 の方法では、検出 5 した発話の音源位置が、既に検出した音源位置のいずれかと一致した場合は更新を行わない。図 7 は第 3 の方法の処理手順を示す。ステップ S 1 で音源数 K の初期値を 0 に設定し、ステップ S 2 で状態検出部 14 により定期的に発話検出動作を実行し、発話状態を検出するとステップ S 3 で音源位置検出部 15 で音源位置の検出を実行する。ステップ S 4 で検出音源位置が既に検出された音源位置のいずれかと一致するか判定し、一致するものがあれば更新処理を行わずステップ S 2 に戻る。ステップ S 4 で既に検出した位置と一致するものがない場合、即ち、前回までと異なる位置にいずれかの音源 9_k が移動したか、又は新たな音源が加わった場合は、ステップ S 5 で K を 1 だけ増加し、ステップ S 6 でその音源に対応する共分散行列 $R_{SkSk}(\omega)$ を共分散行列算出部 17 で新たに計算し、ステップ S 7 で共分散行列記憶 10 部 18 の対応する新たな領域 MA_k に保存すると共に、ステップ S 8 でその共分散行列を使って収音レベル推定部 19 により収音レベルを推定し、ステップ S 9 で全ての共分散行列と推定収音レベルを使ってフィルタ係数算出部 21 により更新したフィルタ係数の算出を実行し、ステップ S 10 でフィルタ $12_1 \sim 12_M$ に更新したフィルタ係数を設定し、ステップ S 2 に戻る。

20 以上示したように、本発明では、複数マイクロホンの受音信号から、音源位置を推定し、各音源に対し受音信号の共分散行列を求め、音源位置ごとに音量を調整するフィルタ係数を求め、それらのフィルタ係数でマイクロホン受音信号をフィルタリングすることで、話者位置ごとに音量を調整した出力信号を得ることができる。

25 図 1 の実施例では、音源位置検出部 15 は各音源 9_k の座標位置を推定する場合について説明したが、音源方向、即ち、マイクロホン $11_1 \sim 11_M$ の配列に対する各音源 9_k の角度位置を求めてよい。音源方向の推定方法は、例えば田中、金田、小島による「音源方向推定方法の室内残響下での性能評価」、日本音響学会誌 Vol. 50, No. 7, 1994, pp. 540-548 に示されている。要するに、各音源に対応して受音信号の共分散行列を求め、記憶しておけばよい。

第2実施例

図8は、本発明の第2の実施例である収音装置の機能ブロック図である。

本実施例の収音装置は、マイクロホン $11_1 \sim 11_M$ と、フィルタ $12_1 \sim 12_M$ と、加算器13と、状態判定部14と、音源位置検出部15と、周波数領域変換部16と、共分散行列算出部17と、共分散行列記憶部18と、収音レベル推定部19と、フィルタ係数算出部21により構成される。

本実施例は、本発明の第1の実施例である収音装置の収音レベル調整に雑音抑圧を加えたものである。

まず、状態判定部14では、マイクロホン $11_1 \sim 11_M$ で受音したマイクロホン受音信号のパワーから、発話区間および雑音区間を検出する。状態判定部14は図9に示すように、図2の状態判定部14の構成に更に雑音判定部14Fが追加された構成となっている。例えば、第1実施例と同様に、それぞれのマイクロホン受音信号について、短時間平均パワー P_{avS} と、長時間平均パワー P_{avL} を短時間平均パワー算出部14Bと長時間平均パワー算出部14Cで求め、割り算部14Dで短時間平均パワーと長時間平均パワーの比 $R_p = P_{avS}/P_{avL}$ が求められ、発話判定部14Eで発話の閾値 P_{thU} と比較され、それを超えていれば発話区間と判定される。雑音判定部14Fはパワー比 R_p を雑音閾値 P_{thN} と比較し、それより小であれば雑音区間であると判定する。

発話判定部14Eの判定結果が発話区間であった場合に、音源位置検出部15は、本発明の第1の実施例と同様にして、音源位置を検出する。

次に、周波数領域変換部16は各音源 9_k の発音区間及び雑音区間においてそれぞれのマイクロホン $11_1 \sim 11_M$ からの受音信号を周波数領域信号に変換して共分散行列演算部17に与える。共分散行列算出部17は、本発明の第1の実施例と同様にして、各音源 9_k に対する周波数領域受音信号の共分散行列 $R_{SkSk}(\omega)$ を算出する。更に、雑音区間における周波数領域受音信号の共分散行列 $R_{NN}(\omega)$ を算出する。

共分散行列記憶部18では、音源位置検出部15の検出結果および状態判定部15の判定結果に基づき、音源 $9_1, \dots, 9_K$ に対する共分散行列 $R_{SkSk}(\omega)$ と雑音区間の共分散行列 $R_{NN}(\omega)$ を領域 $MA_1, \dots, MA_K, MA_{K+1}$ に保存する。

収音レベル推定部19は、各音源の収音レベル P_{Sk} を、本発明の第1の実施例と

同様にして推定する。

次に、フィルタ係数算出部 21 では、各音源 9_k から発せられた音を所望の音量で収音し、かつ雑音を抑圧するためのフィルタ係数を算出する。まず、雑音抑圧の条件を求める。雑音区間のマイクロホン受音信号の周波数領域変換信号を $X_{N,1}(\omega)$ ~ $X_{N,M}(\omega)$ とする。雑音区間のマイクロホン受音信号 $X_{N,1}(\omega)$ ~ $X_{N,M}(\omega)$ がフィルタ 12₁ ~ 12_M および加算器 13 を通過後に 0 となれば、雑音を抑圧できているので、雑音抑圧のための条件式は、式(17)となる。

$$(X_{N,1}(\omega), \dots, X_{N,M}(\omega))H(\omega) = 0 \quad (17)$$

この式(17)と、本発明の第1の実施例で示した収音レベル調整のための条件式(15)の両方を満たせば、収音レベル調整と雑音抑圧の両方を実現できる。

次に、式(15)、(17)の条件をフィルタ係数行列 $H(\omega)$ について最小二乗解で解けば、次式(18)

$$H(\omega) = \left\{ \sum_{k=1}^K C_{Sk} R_{SkSk}(\omega) + C_N R_{NN}(\omega) \right\}^{-1} \sum_{k=1}^K C_{Sk} \sqrt{\frac{P_{opt}}{P_{Sk}}} R_{SkSk}(\omega) A_k(\omega) \quad (18)$$

となる。 C_N は雑音消去量に対する重みの定数であり、値が大きくなるほど雑音の消去量が増加する。ただし、 C_N を大きくすることにより音源位置に対する感度拘束が弱くなり、収音信号周波数特性の劣化が大きくなるため、通常、 C_N は 0.1 ~ 10.0 程度の適当な値に設定する。その他の記号の意味については、第1の実施例と同様である。

次に、式(18)により求められた、フィルタ係数は、フィルタ 12₁ ~ 12_M に設定され、マイクロホン受音信号をそれぞれフィルタリングする。フィルタリング後の信号は、加算器 13 で加算され、出力信号として出力される。

以上、本発明の第2の実施例は、本発明の第1の実施例における収音レベル調整の効果に加え、雑音抑圧を行うことが可能である。

これら以外の部分に関しては、本発明の第1の実施例と同じであるので、説明を省略する。

第3実施例

図10は、本発明の第3の実施例である収音装置の機能ブロック図である。

本実施例の収音装置は、スピーカ 22 と、マイクロホン 11₁～11_M と、フィルタ 12₁～12_M、23 と、加算器 13 と、状態判定部 14 と、音源位置検出部 15 と、周波数領域変換部 16 と、共分散行列算出部 17 と、共分散行列記憶部 18 と、収音レベル推定部 19 と、フィルタ係数算出部 21 により構成される。

5 本実施例は、図 8 に示した本発明の第 2 の実施例である収音装置に受話信号を再生するスピーカ 22 と、受話信号をフィルタリングするフィルタ 23 とを追加し、第 2 実施例による収音レベル調整および雑音抑圧に加えてマイクロホン 11₁～11_M で収音されるスピーカ再生信号成分である音響エコーを抑圧するようにしたものである。

10 状態判定部 14 は、図 11 に示すように、図 4 の状態判定部 14 の構成に、更に受話信号の短時間平均パワー P'_{avS} 及び長時間平均パワー P'_{avL} を算出する短時間平均パワー算出部 14B' と長時間平均パワー算出部 14C' と、これらの比 $R'_p = P'_{avS}/P'_{avL}$ を計算する割り算部 14D' と、その比 R'_p を予め決めた受話信号の閾値 R_{thR} と比較し、それより大であれば受話区間であると判定する受話判定部 14G と、発話判定部 14E、雑音判定部 14F、受話判定部 14G の判定結果に基づいて状態を決定する状態決定部 14H が追加されている。状態決定部 14H は、受話判定部 14G の判定結果が受話区間と判定した場合は、他の発話判定部 14E、雑音判定部 14F の判定結果にかかわらず、受話区間と決定し、受話判定部 14G が受話区間でないと判定した場合は、図 4 の場合と同様に発話判定部 14E 及び雑音判定部 14F の判定に従って、発話区間又は雑音区間と決定する。

15 状態判定部 14 の判定結果が発話区間であった場合に、音源位置検出部 15 は、本発明の第 1 の実施例と同様にして、音源位置を検出する。

20 次に、周波数領域変換部 16 でマイクロホン受音信号および受話信号をそれぞれ周波数領域信号 $X_1(\omega), \dots, X_M(\omega)$ 及び $Z(\omega)$ に変換し、共分散行列算出部 17 はこれらの周波数領域受音信号及び受話信号の共分散行列を生成にする。マイクロホン受音信号の周波数領域変換信号 $X_1(\omega) \sim X_M(\omega)$ 及び受話信号の周波数領域変換信号 $Z(\omega)$ の共分散行列 $R_{XX}(\omega)$ は、次式（19）により算出される。

$$R_{XX}(\omega) = \begin{pmatrix} Z(\omega) \\ X_1(\omega) \\ \vdots \\ X_M(\omega) \end{pmatrix} (Z(\omega)^* X_1(\omega)^* \cdots X_M(\omega)^*) \quad (19)$$

ただし、*は複素共役を表す。

次に、共分散行列記憶部 1 8 では、音源位置検出部 1 5 の検出結果および状態判定部 1 4 の判定結果に基づき、共分散行列 $R_{XX}(\omega)$ を、発話区間における各音源 9 k 5 に対する受音信号及び受話信号の共分散行列 $R_{SKSK}(\omega)$ と、雑音区間における受音信号及び受話信号の共分散行列 $R_{NN}(\omega)$ と、受話区間における受音信号と受話信号の共分散行列 $R_{EE}(\omega)$ としてそれぞれ領域 $MA_1, \dots, MA_K, MA_{K+1}, MA_{K+2}$ に保存する。

10 収音レベル推定部 1 9 は、各音源 9 k の収音レベル P_{Sk} を、各音源に対する共分散行列 $R_{S1S1}, \dots, R_{SKSK}$ と、各音源ごとの M+1 要素からなる予め決めた重み付けミキシングベクトル $A_1(\omega), \dots, A_K(\omega)$ に基づく次式 (20) により求める。

$$P_{Sk} = \frac{1}{W} \sum_{\omega=0}^W A_k(\omega)^H R_{SkSk}(\omega) A_k(\omega) \quad (20)$$

次に、フィルタ係数算出部 2 1 では、各音源から発せられた音を所望の音量で収音するためのフィルタ係数を計算する。まず、各マイクロホンに接続されたのフィルタ 1 2 1 ~ 1 2 M のフィルタ係数を周波数領域に変換したものを $H_1(\omega) \sim H_M(\omega)$ 15 とし、受話信号をフィルタリングするフィルタ 2 3 のフィルタ係数を周波数領域に変換したものを $F(\omega)$ とする。次に、これらのフィルタ係数を式 (21) により行列としたものを $H(\omega)$ とする。

$$H(\omega) = \begin{pmatrix} F(\omega) \\ H_1(\omega) \\ \vdots \\ H_M(\omega) \end{pmatrix} \quad (21)$$

また、受話区間のマイクロホン受音信号の周波数領域変換信号を $X_{E,1}(\omega) \sim X_{E,M}(\omega)$ 、受話信号の周波数領域変換信号を $Z_E(\omega)$ とし、雑音区間のマイクロホン受音信号の周波数領域変換信号を $X_{N,1}(\omega) \sim X_{N,M}(\omega)$ 、受話信号の周波数領域変換信号を $Z_N(\omega)$ とし、k 番目の音源 9 k が発音している発話区間のマイクロホン受音信号の周波数

領域変換信号を $X_{Sk,1}(\omega) \sim X_{Sk,M}(\omega)$ 、受話信号の周波数領域変換信号を $Z_{Sk}(\omega)$ とする。

ここで、フィルタ係数行列 $H(\omega)$ に要求される条件は、マイクロホン受音信号および送話信号をフィルタ係数行列 $H(\omega)$ でそれぞれフィルタリングし、フィルタリング後の信号を加算したときに、音響エコー信号と雑音信号が消去され、送話音声信号のみが所望のレベルで送話されることである。
5

従って、受話区間および雑音区間の信号に対しては、フィルタリングおよび加算後の信号が 0 となる次式 (22) 及び (23)

$$(Z_E(\omega) \ X_{E,1}(\omega) \ \dots \ X_{E,M}(\omega))H(\omega) = 0 \quad (22)$$

$$10 \quad (Z_N(\omega) \ X_{N,1}(\omega) \ \dots \ X_{N,M}(\omega))H(\omega) = 0 \quad (23)$$

が理想条件となり、発話区間の信号に対しては、フィルタリングおよび加算後の信号がマイクロホン受音信号および受話信号に予め決めた $M+1$ 要素の重み付けミキシングベクトル $A_k(\omega)$ と所望の利得を乗じた信号となる次式

$$(Z_{Sk}(\omega) \ X_{Sk,1}(\omega) \ \dots \ X_{Sk,M}(\omega))H(\omega) = \sqrt{\frac{P_{opt}}{P_{Sk}}} (Z_{Sk}(\omega) \ X_{Sk,1}(\omega) \ \dots \ X_{Sk,M}(\omega))A_k(\omega)$$

$$15 \quad (24)$$

が理想条件となる。重み付けミキシングベクトル $A_k(\omega) = (a_0(\omega), a_{k1}(\omega), \dots, a_{kM}(\omega))$ の要素 $a_0(\omega)$ は受話信号に対する重み係数を表し、通常は $a_0(\omega)=0$ と設定する。

次に、式 (22) ~ (24) の条件をフィルタ係数行列 $H(\omega)$ について最小二乗解で解けば、次式

$$20 \quad H(\omega) = \left\{ \sum_{k=1}^K C_{Sk} R_{SkSk}(\omega) + C_N R_{NN}(\omega) + C_E R_{EE}(\omega) \right\}^{-1} \sum_{k=1}^K C_{Sk} \sqrt{\frac{P_{opt}}{P_{Sk}}} R_{SkSk}(\omega) A_k(\omega) \quad (25)$$

となる。 C_E は音響エコー消去量に対する重みの定数であり、値が大きくなるほど音響エコーの消去量が増加する。ただし、 C_E を大きくすることにより収音信号の周波数特性の劣化が大きくなり、雑音抑圧特性が低下する。このため、通常、 C_E は 0.1 ~ 10.0 程度の適当な値に設定する。その他の記号の意味については、第 2 の実施例と同様である。
25

このようにして、音量を調整し、雑音を抑圧するようにフィルタ係数を決定する

ことができる。

次に、式(2 5)により求められた、フィルタ係数は、フィルタ $12_1 \sim 12_M$ 、
3に設定され、マイクロホン受音信号と受話信号をそれぞれフィルタリングする。

フィルタリング後の信号は、加算器 13で加算され、出力信号として出力される。

5 これら以外の部分に関しては、本発明の第2の実施例と同じであるので、説明を省略する。

以上、本発明の第3の実施例は、本発明の第2の実施例における収音レベル調整、
雑音抑圧の効果に加え、音響エコー抑圧を行うことが可能である。第3実施例では
第2実施例において音響エコー抑圧機能を付加した場合を示したが、第1実施例に
10 対し更に音響エコー抑圧機能を付加してもよい。その場合は、図10の状態判定部
14の詳細を示す図11において雑音判定部14Fを除去し、図10の共分散行列
算出部17は雑音区間における共分散行列 $R_{NN}(\omega)$ の算出は行わない。従って、フ
ィルタ係数算出部21におけるフィルタ係数の算出は次式

$$H(\omega) = \left\{ \sum_{k=1}^K C_{Sk} R_{SkSk}(\omega) + C_E R_{EE}(\omega) \right\}^{-1} \sum_{k=1}^K C_{Sk} \sqrt{\frac{P_{opt}}{P_{Sk}}} R_{SkSk}(\omega) A_k(\omega) \quad (26)$$

15 により行えばよいことは前述の説明から明らかである。

第4実施例

図10の第3実施例は、第2実施例の収音レベル調整機能と雑音抑圧機能に更に
音響エコー消去機能を付加した実施例として示したが、雑音抑圧と音響エコー消去
機能のみを有する収音装置として構成してもよい。その構成例を図12に示す。

図12に示すように、この実施例の構成は図10の構成から音源位置検出部15
と収音レベル推定部19を除去し、共分散行列算出部17は送話信号の共分散行列
 $R_{SS}(\omega)$ と、受話信号の共分散行列 $R_{EE}(\omega)$ と、雑音信号の共分散行列 $R_{NN}(\omega)$ を演算
し、共分散行列記憶部18の記憶領域 MA_S, MA_E, MA_N にそれぞれ保存する構成と
25 されている。また、この音響エコー消去機能を実現するには、マイクロホンは少な
くとも1つあればよいが、ここではM個のマイクロホンを使用する例を示している。

状態判定部14は、図10の実施例と同様にマイクロホン $12_1 \sim 12_M$ で受音
した受音信号、及び受話信号から、発話区間、受話区間、及び雑音区間を判定し、

その具体的構成及び動作は図11に示したものと同様である。受音信号及び受話信号は周波数領域変換部16で周波数領域の受音信号 $X_1(\omega) \sim X_M(\omega)$ と周波数領域受話信号 $Z(\omega)$ に変換され、共分散行列算出部17に与えられる。

次に、共分散行列算出部17はそれらの周波数領域受音信号及び受話信号の共分散行列を生成する。マイクロホン受音信号の周波数領域変換信号 $X_1(\omega) \sim X_M(\omega)$ 及び受話信号の周波数領域変換信号 $Z(\omega)$ の共分散行列 $R_{XX}(\omega)$ は、次式(27)により算出される。

$$R_{XX}(\omega) = \begin{pmatrix} Z(\omega) \\ X_1(\omega) \\ \vdots \\ X_M(\omega) \end{pmatrix} (Z(\omega) * X_1(\omega) * \dots * X_M(\omega) *) \quad (27)$$

ただし、*は複素共役を表す。

10 次に、共分散行列記憶部18では、状態判定部14の判定結果に基づき、共分散行列 $R_{XX}(\omega)$ を、受話区間における受音信号及び受話信号の共分散行列 $R_{SS}(\omega)$ と、雑音区間における受音信号及び受話信号の共分散行列 $R_{NN}(\omega)$ と、発話区間における受音信号と受話信号の共分散行列 $R_{EE}(\omega)$ としてそれぞれ領域 MA_S, MA_N, MA_E に保存する。

15 次に、フィルタ係数算出部21では、音源から発せられた送話音を収音し、音響エコー及び雑音を抑圧するためのフィルタ係数を計算する。まず、マイクロホン $1_1 \sim 1_1 M$ に接続されたのフィルタ $1_2_1 \sim 1_2 M$ のフィルタ係数を周波数領域に変換したものを $H_1(\omega) \sim H_M(\omega)$ とし、受話信号をフィルタリングするフィルタ 2_3 のフィルタ係数を周波数領域に変換したものを $F(\omega)$ とする。次に、これらのフィルタ係数を式(28)により行列としたものを $H(\omega)$ とする。

$$H(\omega) = \begin{pmatrix} F(\omega) \\ H_1(\omega) \\ \vdots \\ H_M(\omega) \end{pmatrix} \quad (28)$$

また、受話区間のマイクロホン受音信号の周波数領域変換信号を $X_{E,1}(\omega) \sim X_{E,M}(\omega)$ 、受話信号の周波数領域変換信号を $Z_E(\omega)$ とし、雑音区間のマイクロホン受音信号の

周波数領域変換信号を $X_{N,1}(\omega) \sim X_{N,M}(\omega)$ 、受話信号の周波数領域変換信号を $Z_N(\omega)$ とする。発話区間のマイクロホン受音信号の周波数領域変換信号を $X_{S,1}(\omega) \sim X_{S,M}(\omega)$ 、受話信号の周波数領域変換信号を $Z_S(\omega)$ とする。

ここで、フィルタ係数行列 $H(\omega)$ に要求される条件は、マイクロホン受音信号および送話信号をフィルタ係数行列 $H(\omega)$ でそれぞれフィルタリングし、フィルタリング後の信号を加算したときに、音響エコー信号と雑音信号が消去され、送話音声信号のみが送話されることである。

従って、受話区間および雑音区間の信号に対しては、フィルタリングおよび加算後の信号が 0 となる次式 (29) 及び (30)

$$(Z_E(\omega) \ X_{E,1}(\omega) \ \dots \ X_{E,M}(\omega))H(\omega) = 0 \quad (29)$$

$$(Z_N(\omega) \ X_{N,1}(\omega) \ \dots \ X_{N,M}(\omega))H(\omega) = 0 \quad (30)$$

が理想条件となり、発話区間の信号に対しては、フィルタリングおよび加算後の信号がマイクロホン受音信号および受話信号に予め決めた $M+1$ 要素の重み付けミキシングベクトル $A(\omega)$ を乗じた信号となる次式

$$(Z_S(\omega) \ X_{S,1}(\omega) \ \dots \ X_{S,M}(\omega))H(\omega) = (Z_S(\omega) \ X_{S,1}(\omega) \ \dots \ X_{S,M}(\omega))A(\omega) \quad (31)$$

が理想条件となる。重み付けミキシングベクトル $A(\omega) = (a_0(\omega), a_{k1}(\omega), \dots, a_{kM}(\omega))$ の第 1 要素 $a_0(\omega)$ は受話信号に対する重み係数を表し、通常は $a_0(\omega)=0$ に設定する。

次に、式 (29) ~ (31) の条件をフィルタ係数行列 $H(\omega)$ について最小二乗解で解けば、次式

$$H(\omega) = \{R_{SS}(\omega) + C_N R_{NN}(\omega) + C_E R_{EE}(\omega)\}^{-1} R_{SS}(\omega) A(\omega) \quad (32)$$

となる。 C_E は音響エコー消去量に対する重みの定数であり、値が大きくなるほど音響エコーの消去量が増加する。ただし、 C_E を大きくすることにより収音信号の周波数特性の劣化が大きくなり、雑音抑圧特性が低下する。このため、通常、 C_E は 0.1 ~ 10.0 程度の適当な値に設定する。その他の記号の意味については、第 2 の実施例と同様である。

このようにして、音量を調整し、雑音を抑圧するようにフィルタ係数を決定することができる。

次に、式 (32) により求められたフィルタ係数は、フィルタ $1\ 2_1 \sim 1\ 2_M$ 及び $2\ 3$ に設定され、マイクロホン受音信号と受話信号をそれぞれフィルタリングする。

フィルタリング後の信号は、加算器 1 3 で加算され、出力信号として出力される。これら以外の部分に関しては、本発明の第 2 の実施例と同じであるので、説明を省略する。

以上、本発明の第 4 の実施例は、雑音抑圧の効果に加え、音響エコー抑圧を行う
5 ことが可能である。

第 5 実施例

図 1 3 に第 5 実施例を示す。第 5 実施例は、図 1 2 に示した第 4 実施例において、
10 発話区間に音源位置を検出し、音源位置ごとに共分散行列を求めこれらを保存して
おき、雑音区間では雑音に対する共分散行列を求めて保存しておく。次にこれら保
15 存されている共分散行列を用いて雑音と音響エコーを抑圧するフィルタ係数を求
める。これらフィルタ係数でマイクロホン受音信号と受話信号をフィルタリングする
ことで、雑音と音響エコーを抑圧した送話信号を得る。

この第 5 実施例の構成は、図 1 0 に示した第 3 実施例において収音レベル推定部
15 1 9 を除去した構成と同等である。

状態判定部 1 4 では、第 3 実施例と同様に発話区間、受話区間及び雑音区間を検
出する。状態判定部 1 4 の判定結果が発話区間であった場合に、音源位置検出部 1
5 は各音源 9_k の位置を推定する。音源位置の推定方法は、図 1 で示した第 1 実施
例と同様であるので説明を省略する。

20 次に、周波数領域変換部 1 6 で受音信号及び受話信号は周波数領域信号に変換さ
れ、共分散行列算出部 1 7 に与えられる。

25 共分散行列算出部 1 7 は、各音源 9_k に対する受音信号と受話信号の共分散行列
 $R_{S1S1}(\omega) \sim R_{SKSK}(\omega)$ と、受話区間の共分散行列 $R_{EE}(\omega)$ と、雑音区間の共分散行列
 $R_{NN}(\omega)$ を算出する。共分散行列記憶部 1 8 は状態判定部 1 4 の判定結果及び音源位
置検出部 1 5 の位置検出結果に基づいて、共分散行列 $R_{S1S1}(\omega) \sim R_{SKSK}(\omega), R_{EE}(\omega),$
 $R_{NN}(\omega)$ を対応する領域 $MA_1 \sim MA_K, MA_{K+1}, MA_{K+2}$ に保存する。

フィルタ係数算出部 2 1 は、送話音声が収音されると、音響エコー及び雑音を抑
圧するフィルタ係数を計算する。第 3 実施例と同様に、フィルタ係数行列 $H(\omega)$ に
についての条件式を最小二乗解で解けば、次式

$$H(\omega) = \left\{ \sum_{k=1}^K C_{Sk} R_{SkSk}(\omega) + C_N R_{NN}(\omega) + C_E R_{EE}(\omega) \right\}^{-1} \sum_{k=1}^K C_{Sk} R_{SkSk}(\omega) A_k(\omega) \quad (33)$$

となる。ただし、 $C_{S1} \sim C_{SK}$ は各音源に対する感度拘束の重み定数、 C_E は音響エコー消去量に対する重みの定数、 C_N は雑音消去量に対する重み定数である。

このようにして求められたフィルタ係数は、フィルタ $12_1 \sim 12_M, 23$ に設定

5 され、マイクロホン受音信号と受話信号をそれぞれフィルタリングする。フィルタリング後の信号は、加算器 13 で加算され、出力信号として出力される。これら以外の部分に関しては、本発明の第 3 の実施例と同じであるので、説明を省略する。この第 5 実施例により、第 4 実施例と同様に音響エコーと雑音を抑圧した送話信号を生成することができる。また、第 5 実施例では、複数音源位置について感度を拘束でき、過去に発話した音源に対しても感度を保持しておくことが可能である。従って、音源位置の移動があっても、過去に発話した音源であれば感度が保持されているので、送話音声の話頭の音質劣化がないという利点がある。

第 6 実施例

15 本発明の第 6 の実施例である収音装置について説明する。

本発明の第 6 の実施例である収音装置では、本発明の第 1 ~ 3, 5 の実施例である収音装置の各音源位置 9_k に対する感度拘束の重み係数 $C_{S1} \sim C_{SK}$ を時間的に変化させる。

音源 $9_1 \sim 9_K$ に対する時間的に変化する感度拘束の重み係数 $C_{S1} \sim C_{SK}$ は、過去 20 に発話された音源ほど小さくする。その第 1 の方法は、既に検出されているそれぞれの音源位置の検出時刻から最も最近に検出された音源位置の検出時刻までの経過時間が長いほど重み係数 C_{Sk} を小さくする。第 2 の方法は、 K 個の音源位置の検出時刻順に重み係数 C_{Sk} を小さくする。

図 14 は上記第 1 の方法を実施する重み係数設定部 $21H$ の機能構成ブロック図 25 を示す。この例では、重み係数設定部 $21H$ は時刻を出力するクロック $21H1$ と、音源位置検出ごとに検出された音源 9_k を表す番号 k をアドレスとしてその検出時刻 t を上書きする時刻記憶部 $21H2$ と、重み係数決定部 $21H3$ とから構成されている。重み係数決定部 $21H3$ は時刻記録部 $21H2$ に記録されている音源位置検出時刻

に基づいて今回検出された番号 $k(t)$ の音源に対し、予め決めた値 C_s を重み係数 C_{Sk} として与え、その他の番号 $k \neq k(t)$ に対してはそれぞれの検出時刻 t_k からの経過時間 $t - t_k$ により $q^{(t-t_k)} C_s$ を重み係数 C_{Sk} として与える。 q は $0 < q \leq 1$ の予め決めた値である。このようにして各音源に対して感度拘束の重み係数 $C_{S1} \sim C_{SK}$ が決定され、
5 乗算器 21A1～21AK に与えられる。

図 15 は上記第 2 の方法を実施する重み係数設定部 21H の機能構成ブロック図を示す。この例ではクロック 21H1 と、時刻記録部 21H2 と、順序判定部 21H4 と、重み係数決定部 21H5 とから構成されている。順序判定部 21H4 は時刻記録部 21H2 に記録されている時刻から音源 $9_1 \sim 9_K$ の位置検出順(最も新しい順) $\{k(t)\} = \{k(1), \dots, k(K)\}$ を判定する。重み係数決定部 21H5 は最も最近の検出音源 $9_{k(1)}$ に対し予め決めた値 C_s を重み係数 $C_{Sk(1)}$ として与える。その他の音源に対し、 $C_{Sk(t+1)} \leftarrow q C_{Sk(t)}$ を $t=1, 2, \dots, K-1$ までそれぞれ計算して重み係数 $C_{Sk(2)}, \dots, C_{Sk(K)}$ を得る。これら $C_{Sk(1)} \sim C_{Sk(K)}$ は $\{k(1), \dots, k(K)\}$ の順に従って並び替えて重み係数 C_{S1}, \dots, C_{SK} として出力される。 q の値は $0 < q < 1$ の予め決めた値である。
10

15 このように各音源位置に対する感度拘束の重みを変化させることにより、過去に発話のあった音源位置に対する感度拘束を減少させることができる。これにより、本発明の第 1～3 の実施例である収音装置に比べ、感度拘束する音源数を抑えることができ、収音レベル調整の性能および雑音、音響エコー抑制性能を向上することができる。

20 これら以外の部分に関しては、本発明の第 1～3、5 等の実施例と同じであるので、説明を省略する。

第 7 実施例

本発明の第 7 の実施例である収音装置について説明する。

25 本発明の第 7 の実施例である収音装置は、本発明の第 1～6 の実施例である収音装置のフィルタ係数算出部 21 において、共分散行列 $R_{xx}(\omega)$ を白色化することを特徴とする。図 16 は図 4 のフィルタ係数算出部 21 中に破線で示す白色化部 21J1～21JK の 1 つを代表してその機能構成を示す。この白色化部 21J は対角行列算出部 21JA と、重み付け部 21JB と逆演算部 21JC と、乗算部 21JD とにより構成されてい

る。対角行列算出部 21JA は与えられた共分散行列 $\mathbf{R}_{xx}(\omega)$ の対角行列 $\text{diag}(\mathbf{R}_{xx}(\omega))$ を生成する。重み付け部 21JB は予め決めた任意の M 行または M+1 行の行列 \mathbf{D} に基づく重みにより、次式を計算して対角行列に重みを与える。

$$\mathbf{D}^T \text{diag}(\mathbf{R}_{xx}(\omega)) \mathbf{D} \quad (34)$$

5 逆演算部 21JC は式 (34) の逆数

$$1/\{\mathbf{D}^T \text{diag}(\mathbf{R}_{xx}(\omega)) \mathbf{D}\} \quad (35)$$

を計算する。ただし、^Tは行列の転置を表す。この逆演算部 21JC の演算結果は入力されたそれぞれの共分散行列 $\mathbf{R}_{xx}(\omega)$ と乗算部 21JD で乗算され、白色化された共分散行列が得られる。

10 共分散行列の白色化により、フィルタ係数算出部 21 において求められるフィルタ係数は、送話信号、受話信号、雑音信号のスペクトルの変化に対して、変化しなくなる。これにより、スペクトル変化による収音レベル調整性能、および音響エコー、雑音の抑圧性能の変化がなくなり、定常的な収音レベル調整と音響エコー、雑音抑圧を得ることができる。

15 これら以外の部分に関しては、本発明の第 1 ~ 4 の実施例と同じであるので、説明を省略する。

第 8 実施例

本発明の第 8 の実施例である収音装置について説明する。

20 本発明の第 8 の実施例である収音装置は、本発明の第 1 ~ 7 の実施例である収音装置の共分散行列記憶部 18 が、既に記憶されている共分散行列と、共分散行列算出部 17 により新たに算出された共分散行列を平均して、現在の共分散行列として記憶することを特徴とする。

25 共分散行列の平均は、例えば以下の方法により行われる。既に記憶されている共分散行列を $\mathbf{R}_{xx,\text{old}}(\omega)$ とし、共分散行列算出部 17 により新たに算出された共分散行列を $\mathbf{R}_{xx,\text{new}}(\omega)$ とし、次式

$$\mathbf{R}_{xx}(\omega) = (1-p)\mathbf{R}_{xx,\text{new}}(\omega) + p\mathbf{R}_{xx,\text{old}}(\omega) \quad (36)$$

により平均された共分散行列 $\mathbf{R}_{xx}(\omega)$ を求める。ただし、p は、平均の時定数を決定する定数であり、 $0 \leq p < 1$ の値をとる。

図17は共分散行列記憶部18とその中に設けられた平均化部18Aの機能構成を示す。平均化部18Aは乗算器18A1と、加算器18A2と、乗算器18A3とから構成されている。共分散行列算出部17により算出された音源 9_k に対応する共分散行列 $R_{SkSk}(\omega)$ は新共分散行列 $R_{SkSk,new}(\omega)$ として乗算器18A1に与えられ、 $(1-p)$ と乗算され、その乗算結果は加算器18A2に与えられる。一方、音源 9_k に対応する共分散行列 $R_{SkSk}(\omega)$ が記憶領域18Bから読み出され、旧共分散行列 $R_{SkSk,old}(\omega)$ として乗算器18A3に与えられ、定数 p と乗算される。その乗算結果は加算器18A2により乗算器18A1の出力 $(1-p)R_{SkSk,new}(\omega)$ と加算され、得られた平均共分散行列 $R_{SkSk}(\omega)$ が音源 9_k に対応する記憶領域に上書きされる。

以上のように、共分散行列を平均して記憶することにより、平均前より回路雑音などの外乱の影響を減らし、正確な共分散行列を求めることができ、従って、取音レベル調整、雑音抑圧、あるいは音響エコー抑圧の性能を向上させたフィルタ係数を決定することができる。

これら以外の部分に関しては、本発明の第1～5の実施例と同じであるので、説明を省略する。

なお、本発明は専用のハードウェアにより実現することもできるし、その機能を実現するためのプログラムを、コンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータに読み込ませ、実行することもできる。コンピュータ読み取り可能な記録媒体とは、フロッピーディスク、光磁気ディスク、CD-ROM、DVD-ROM、不揮発性半導体メモリ、内蔵または外付けハードディスク、等の記憶装置をさす。更に、コンピュータ読み取り可能な記録媒体は、インターネットを介してプログラムを送信する場合のように、短時間の間、動的にプログラムを保持するもの（伝送媒体もしくは伝送は）、その場合のサーバとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含む。

発明の効果

次に、本発明による収音装置の第1実施例の有効性を示すために、図18A、1

8 B に一辺が 20 cm の正方形のそれぞれの角にマイクロホンを配置した場合のシミュレーション結果を示す。シミュレーション条件は、マイクロホン数：4、信号対雑音比：約 20dB、部屋の残響時間：300ms、話者数：2（話者 A は 4 つのマイクロホンの中心から正方形の一辺と直角方向に 50cm の位置、話者 B は上記中心から 5 話者 A と 90° 成す方向に 200cm の位置）である。これらの条件のもとで、話者 A と話者 B が交互に発音した場合のマイクロホン受音信号波形を図 18 A に示す。話者 A と話者 B の音声波形を比較すると、話者 B の音声の振幅が低くなっていることが分かる。これに対し、本発明により処理を行った後の波形を図 18 B に示す。話者 A と話者 B の音声波形の振幅の大きさがほぼ同じとなっており、収音レベル調整 10 の効果が確認できる。

図 19 は図 10 に示した第 3 実施例によるシミュレーション結果を示す。シミュレーション条件は、マイクロホン数 M : 4、処理前の送話信号対雑音比 : 20 dB、送話信号対音響エコー信号比 : -10 dB、部屋の残響時間 : 300 msec である。この条件で、送話と受話を交互に繰り返し、そのときの送話信号レベルを図 19 に示す。行 A が処理前の送話信号レベルであり、行 B が第 3 実施例により処理した後の送話信号レベルである。以上の結果より、第 3 実施例は音響エコー信号を約 40 dB、雑音信号を約 15 dB 低減していることがわかり、この発明の実施例が有効であることが確認できる。

以上説明したように、本発明の第 1 実施例によれば、複数のマイクロホンで受音 20 した信号から音源位置を検出し、音源位置ごとの発話区間の共分散行列に基づきフィルタ係数を算出し、これらのフィルタ係数で、マイクロホン受音信号をフィルタリングして加算することにより、音量を音源位置ごとに調整した送話信号を得ることができる。

この発明の第 2 実施例によれば、第 1 実施例において発話区間の共分散行列に加え更に雑音区間の共分散行列を使用してフィルタ係数を決定することにより、収音 25 レベル調整に加えて雑音抑圧が可能になる。

この発明の第 3 実施例によれば、第 1 又は第 2 実施例において、発話区間の共分散行列に加えて更に受話区間における共分散行列を使用してフィルタ係数を決定することにより、音響エコーの抑圧が可能となる。

この発明の第4実施例によれば、発話区間の共分散行列と受話区間の共分散行列を使ってフィルタ係数を決めることにより、受話信号がスピーカで再生され、収音された音響エコーを抑圧することができる。

この発明の第5実施例によれば、第4実施例において、発話区間及び受話区間の
5 共分散行列に加え更に雑音区間の共分散行列を使ってフィルタ係数を決めることにより、更に雑音抑圧が可能となる。

この発明の第6実施例によれば、上記第1、2、3及び5の実施例において、フィルタ係数計算時に、より古い発音の共分散行列に対してより小さい感度拘束の重
み係数を与えることにより、収音レベルの調整性能、雑音抑圧性能、あるいは音響
10 エコー抑圧性能を更に向上することができる。

この発明の第7実施例によれば、上記第1乃至第6実施例において、フィルタ係数計算時に、共分散行列を白色化することにより、信号スペクトルの変化による影響の少ない収音レベル調整、雑音抑圧、音響エコー抑圧を可能にする。

この発明の第8実施例によれば、第1乃至第7実施例において、共分散行列の保存時に、対応する領域に既に保存されている共分散行列との重み付き平均共分散行列で更新保存することにより、より正確な共分散行列が得られ、収音レベル調整、雑音抑圧、音響エコー抑圧の性能を向上したフィルタ係数を決めることができる。

請求の範囲

1. 音響空間に配置された複数チャネルのマイクロホンにより各音源から発せられた音を収音する収音方法であって、

5 (a) 上記複数チャネルのマイクロホンで受音された受音信号から、発話区間を判定する発話判定段階を含む状態判定段階と、

(b) 上記発話判定段階において発話区間と判定された場合に、上記受音信号から、各音源位置を検出する音源位置検出段階と、

(c) 上記受音信号を周波数領域の信号に変換する周波数領域変換段階と、

10 (d) 上記周波数領域の受音信号の共分散行列を算出する共分散行列算出段階と、

(e) 上記音源位置検出段階の検出結果に基づき、各音源に対し上記共分散行列を記憶する共分散行列記憶段階と、

(f) 記憶されている上記共分散行列と予め決めた出力レベルに基づき上記複数チャネルのフィルタ係数を算出するフィルタ係数算出段階と、

15 (g) 上記複数チャネルの受音信号をそれぞれ上記複数チャネルのフィルタ係数でフィルタリングするフィルタリング段階と、

(h) 上記複数チャネルのフィルタリング結果を加算して送話信号として出力する加算段階、

20 とを含む収音方法

2. 請求項1記載の収音方法において、更に、各音源に対応して記憶されている共分散行列に基づいて上記音源からの発話に対する収音レベルを推定する収音レベル推定段階を含み、上記フィルタ係数算出段階は上記各音源に対応して記憶された上記共分散行列及び上記推定された収音レベルに基づいて出力レベルが所望のレベルとなるよう、上記複数チャネルのフィルタ係数を算出する段階を含む。

25 3. 請求項2記載の収音方法において、上記状態判定段階は、上記複数チャネルの受音信号から雑音区間を判定する雑音判定段階を含み、

上記共分散行列算出段階は、上記雑音区間と判定された場合、その区間での受音信号の共分散行列を雑音の共分散行列として算出する段階を含み、

上記共分散行列記憶段階は、各音源ごとに対応して上記受音信号の共分散行列を記憶すると共に、上記雑音区間の共分散行列を記憶するよう適応されており、

上記フィルタ係数算出段階は、上記発話区間の各音源に対する記憶された共分散行列と上記雑音区間の記憶された共分散行列に基づいて、各音源に対する収音レベルが各々所望のレベルとなり、かつ雑音を抑圧するように複数チャネルのフィルタ係数を算出するよう適応されている。
5

4. 請求項2記載の収音方法において、上記音響空間に受話信号を再生して出力するスピーカが設けられており、上記状態判定段階は受話信号から受話区間を判定する受話判定段階を含み、

10 上記周波数領域変換段階は受話信号を周波数領域の信号に変換する段階を含み、

上記共分散行列算出段階は、上記発話区間及び上記受話区間のそれぞれにおいて上記共分散行列を周波数領域の上記複数チャネルの受音信号と上記受話信号から算出し、

15 上記共分散行列記憶段階は、上記発話区間の各音源に対応する共分散行列と、上記受話区間の共分散行列を記憶し、

上記フィルタ係数算出段階は、上記受話区間の記憶された共分散行列および上記発話区間の音源ごとに記憶された共分散行列に基づいて、各音源に対する収音レベルが各々所望のレベルとなり、上記受音信号に含まれる上記スピーカによる再生音の信号成分である音響エコーを抑圧するよう複数チャネルのフィルタ係数を算出す
20 る。

5. 請求項1乃至4のいずれか記載の収音方法において、上記音源は2以上のK個あり、上記フィルタ係数算出段階は、K個の音源に対する感度拘束の重み $C_{s1} \sim C_{sk}$ を、過去に発音した音源に対する重みほど小さくしてそれぞれの共分散行列に与えてから上記フィルタ係数を算出する。

25 6. 請求項1乃至4のいずれか記載の収音方法において、上記複数チャネルをMチャネルとすると、上記フィルタ係数算出段階は、各々の共分散行列 $R_{xx}(\omega)$ の対角成分 $\text{diag}(R_{xx}(\omega))$ と任意の M 行または M+1 行の行列 D に基づく重み $1/\{D^H \text{diag}(R_{xx}(\omega)) D\}$ を各々の共分散行列 $R_{xx}(\omega)$ に乘じ白色化してから、上記フィルタ係数の算出を行う。

7. 請求項 1 乃至 4 のいずれか記載の収音方法において、上記共分散行列記憶段階は、過去に記憶されている共分散行列と、上記共分散行列算出段階により新たに算出された共分散行列を平均して、現在の共分散行列として記憶する。

8. 音響空間に配置された複数チャネルのマイクロホンにより各音源から発せられた音を収音する収音装置において、

複数チャネルのマイクロホンの各々で受音された受音信号から、発話区間を判定する発話判定部を含む状態判定部と、

上記発話判定部において発話区間と判定された場合に、上記受音信号から、音源位置を検出する音源位置検出部と、

10 上記受音信号を周波数領域の信号に変換する周波数領域変換部と、

上記周波数領域の複数チャネルの受音信号の共分散行列を各音源ごとに算出する共分散行列算出部と、

上記音源位置検出手段の検出結果に基づき、上記共分散行列を音源ごとに記憶する共分散行列記憶部と、

15 記憶された上記共分散行列を用いて、各音源に対する送話信号レベルが各々所望のレベルとなるよう複数チャネルのフィルタ係数を算出するフィルタ係数算出部と、

上記マイクロホン受音信号を、上記複数チャネルのフィルタ係数で、各々フィルタリングする複数チャネルのフィルタと、

上記複数チャネルのフィルタの出力信号し、送話信号として出力する加算器、
20 とを含む。

9. 請求項 8 記載の収音装置において、各音源に対応して記憶された共分散行列から各音源の収音レベルを推定する収音レベル推定部を更に含み、上記フィルタ係数算出部は上記推定された収音レベルに基づいて各音源に対する送話信号レベルが所定レベルとなるよう対応する共分散行列に重みを与えて上記フィルタ係数を算出す
25 るよう適応されている。

10. 請求項 1 乃至 7 のいずれか記載の収音方法をコンピュータで実行させるための収音プログラム。

11. 受話信号がスピーカから再生される音響空間内で少なくとも 1 つのマイクロホンで少なくとも 1 つの音源からの発音を収音する収音方法であり、

(a) 上記少なくとも 1 チャネルのマイクロホンで受音された受音信号と受話信号から発話区間及び受話区間を判定する状態判定段階と、
(b) 上記受音信号及び上記受話信号を周波数領域の信号に変換する周波数領域変換段階と、
5 (c) 上記周波数領域の受音信号及び受話信号から発話区間での共分散行列及び受話区間での共分散行列を算出する共分散行列算出段階と、
(d) 上記状態判定段階の判定結果に基づいて、上記共分散行列を上記発話区間及び上記受話区間ごとに記憶する共分散行列記憶段階と、
10 (e) 上記発話区間及び上記受話区間での記憶された共分散行列に基づいて受話信号の音響エコーを抑圧するよう上記マイクロホンと同チャネル数の受音信号用フィルタ係数と、 1 チャネルの上記受話信号用フィルタ係数を算出するフィルタ係数算出段階と、
15 (f) 上記受音信号を、上記少なくとも 1 チャネルの受音信号用フィルタ係数でフィルタリングし、上記受話信号を上記受話信号用フィルタ係数でフィルタリングするフィルタリング段階と、
(g) 上記フィルタリングされた信号を加算して送話信号として出力する加算段階、とを含む。

12. 請求項 1 1 記載の収音方法において、上記状態判定段階は、上記受音信号及び上記受話信号から雑音区間を判定する段階を含み、上記共分散行列算出段階は上記雑音区間での共分散行列を算出する段階を含み、上記共分散行列記憶段階は、上記雑音区間での共分散行列を記憶する段階を含み、上記フィルタ係数算出段階は、上記発話区間、上記受話区間及び上記雑音区間での記憶された共分散行列に基づいて上記音響エコーと雑音を抑圧するよう上記少なくとも 1 チャネルの受音信号用フィルタ係数と、上記受話信号用フィルタ係数を算出する。

13. 請求項 1 1 記載の収音方法において、上記マイクロホンは複数チャネルに設けられ、複数音源からの発音を受音し、上記方法は更に、上記状態判定段階により発話区間と判定された場合に、上記複数チャネルのマイクロホンからの受音信号から音源位置を検出する音源位置検出段階含み、上記共分散行列記憶段階は、上記状態判定段階の判定結果及び検出音源位置に基づいて上記共分散行列を音源位置、

受話区間ごとに記憶する。

14. 請求項13の収音方法において、上記フィルタ係数算出段階は、K個の各音源位置に対する感度拘束の重み係数 $C_{s1} \sim C_{sk}$ を、過去に発音した音源に対する重みほど小さくしてそれぞれの音源に対応する上記共分散行列に与えてから上記フィルタ係数を算出する。
5

15. 請求項11乃至14のいずれかの収音方法において、上記複数チャネルは2以上の中チャネルであり、上記フィルタ係数算出段階は、それぞれの $M+1$ 行 $M+1$ 列の共分散行列 $R_{xx}(\omega)$ の対角成分 $\text{diag}(R_{xx}(\omega))$ と任意の $M+1$ 行の行列Dに基づく重み $1/\{D^T \text{diag}(R_{xx}(\omega)) D\}$ をそれぞれの共分散行列 $R_{xx}(\omega)$ に乗じて白色化してから

10 フィルタ係数を算出する。

16. 請求項11乃至14のいずれかの収音方法において、上記共分散行列記憶段階は、既に記憶されている共分散行列と、上記共分散行列算出部により新たに算出された共分散行列を平均して現在の共分散行列として記憶する。

17. 収音装置であり、

15 音源からの送話音を収音して受音信号を出力する少なくとも1チャネルのマイクロホンと、

受話信号を再生出力するスピーカと、

受音信号と受話信号から発話区間及び受話区間を判定する状態判定部と、

上記受話信号及び上記受音信号を周波数領域の信号に変換する周波数領域変換部

20 と、

上記発話区間及び受話区間ごとに上記周波数領域の受音信号と受話信号の共分散行列を算出する共分散行列算出部と、

上記共分散行列を発話区間及び受話区間ごとに記憶する共分散行列記憶部と、

記憶された上記共分散行列に基づいて上記受話信号の音響エコーを抑圧するよう

25 に上記少なくとも1チャネルの受音信号用フィルタ係数と上記受話信号用フィルタ係数を算出するフィルタ係数算出部と、

上記受音信号用フィルタ係数と受話信号用フィルタ係数が設定され、上記受音信号と受話信号をそれぞれフィルタリングする受音信号用フィルタと受話信号用フィルタと、

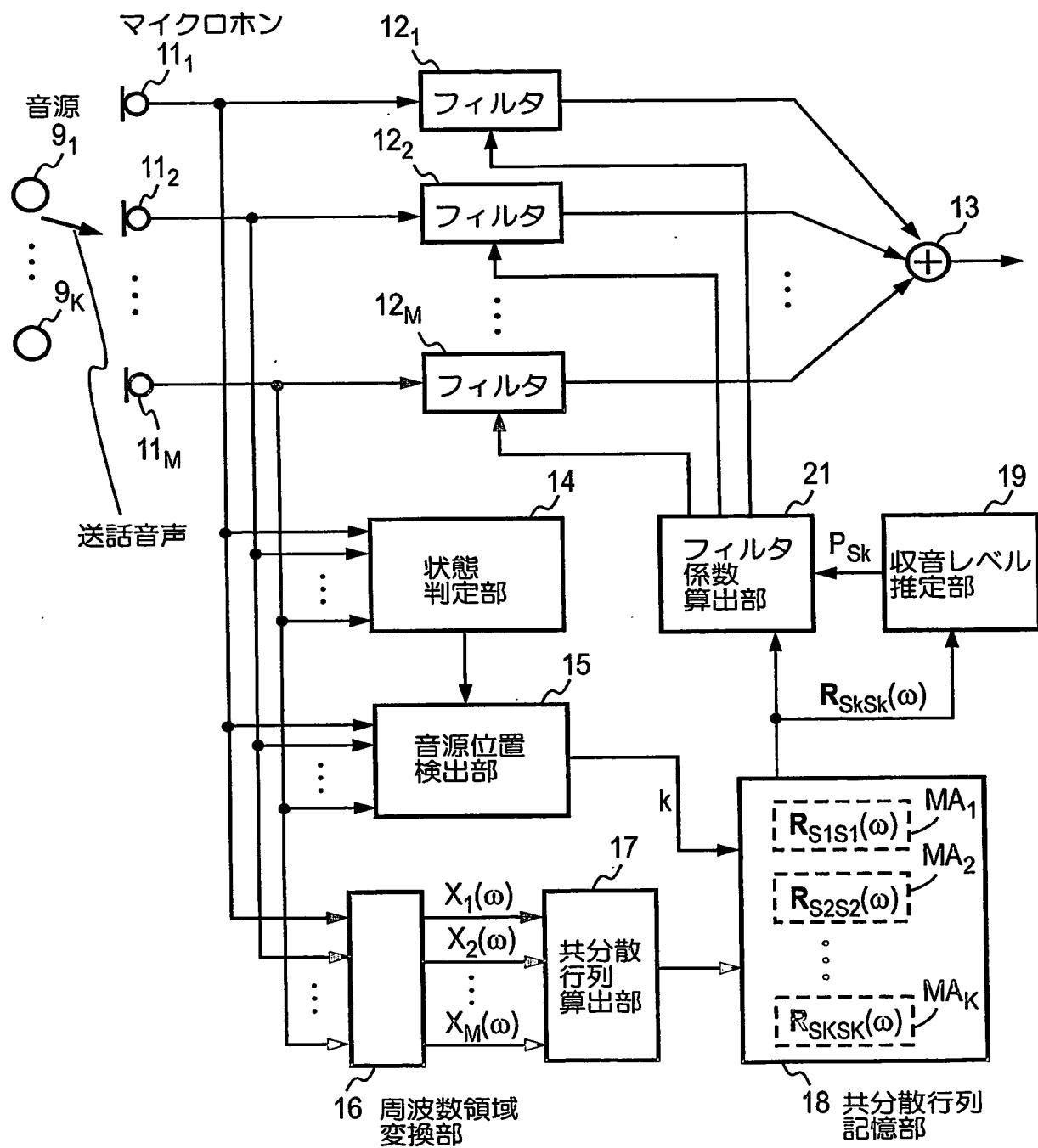
上記受音信号用フィルタの出力と上記受話信号用フィルタの出力を加算して送話信号として出力する加算器、
とを含む。

18. 請求項17記載の収音装置において、上記マイクロホンと上記受音信号用フィルタはそれぞれ複数チャネル設けられており、上記加算器は、上記複数チャネルの受音信号用フィルタの出力と上記受話信号用フィルタの出力を加算して送話信号として出力する。
5

19. 請求項18記載の収音装置において、上記状態判定部は上記受音信号と上記受話信号から雑音区間を判定する雑音判定部を含み、上記共分散行列算出部は上記10
雑音区間の受音信号と受話信号の共分散行列を算出するよう適応されており、上記共分散行列記憶部は上記雑音区間の共分散行列を記憶するよう適応されており、上記
記フィルタ係数算出部は記憶された上記共分散行列に基づいて上記受話信号の音響エコー及び雑音を抑圧するよう上記複数チャネルのフィルタ係数を算出し、上記複数チャネルのフィルタに設定するよう適応されている。

20. 請求項19記載の収音装置において、更に上記複数チャネルの受音信号に基づいてK個の音源の位置を検出する音源位置検出部が設けられ、上記共分散行列算出部は、上記発話区間の共分散行列を各音源ごとに算出するよう適応されており、上記共分散行列記憶部は上記発話区間の共分散行列を各音源ごとに記憶するよう適応されており、上記フィルタ係数算出部はそれぞれの音源に対し過去に発話した音源ほど小さな感度拘束の重み $C_{s1} \sim C_{sK}$ を小さくしてそれぞれの共分散行列に与えて上記フィルタ係数を算出する手段を含む。
15
20

21. 請求項11乃至16のいずれか記載の収音方法をコンピュータで実行させるための収音プログラム。



1

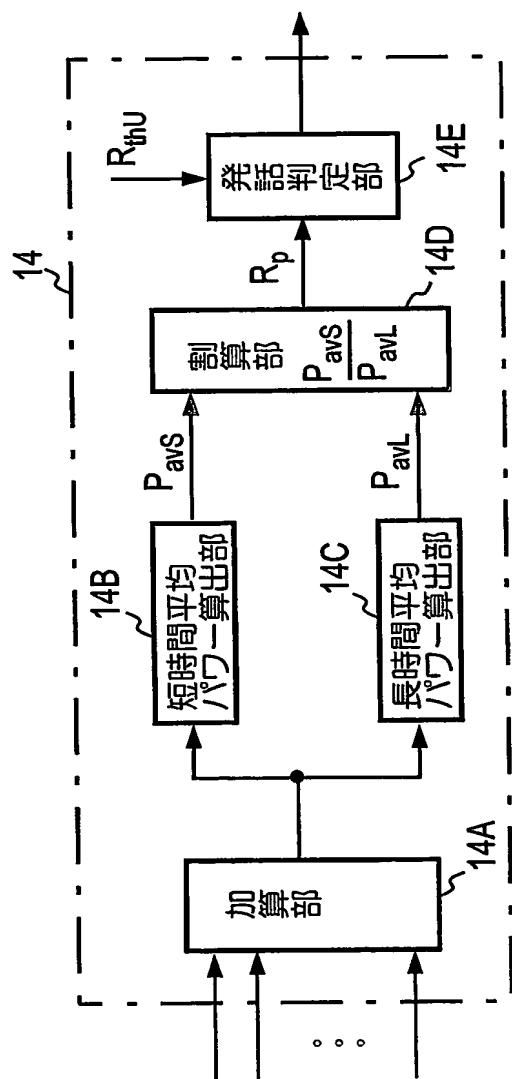


図2

3/18

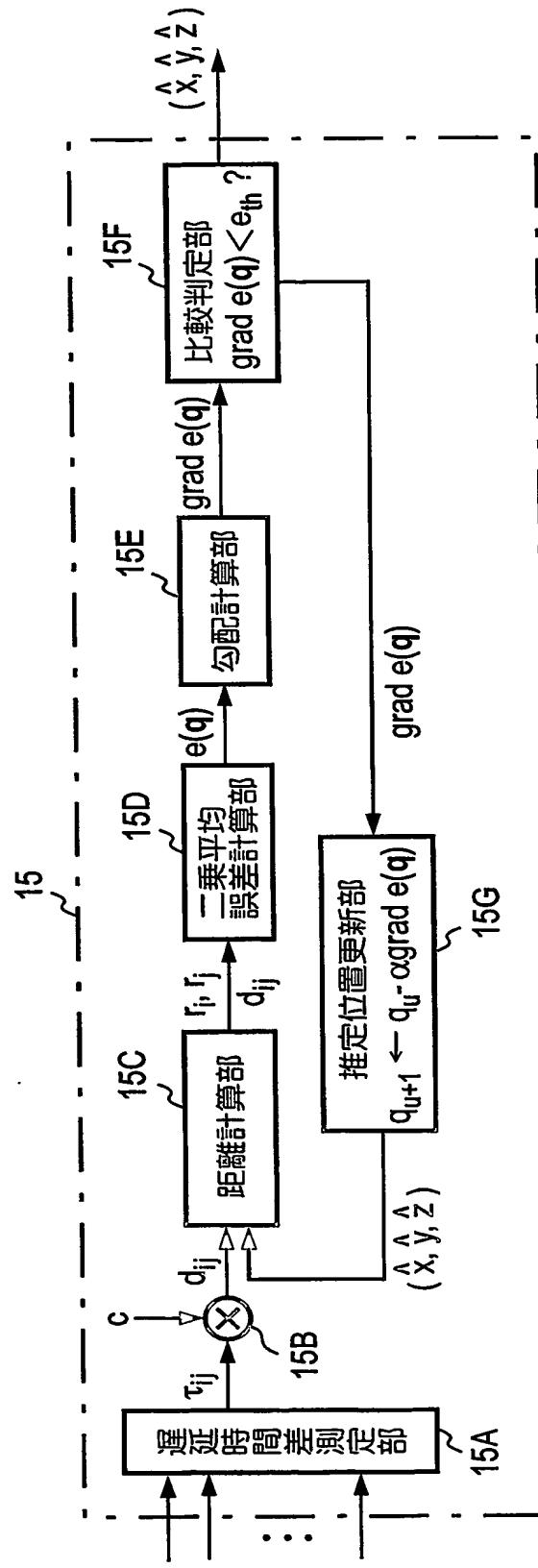


図3

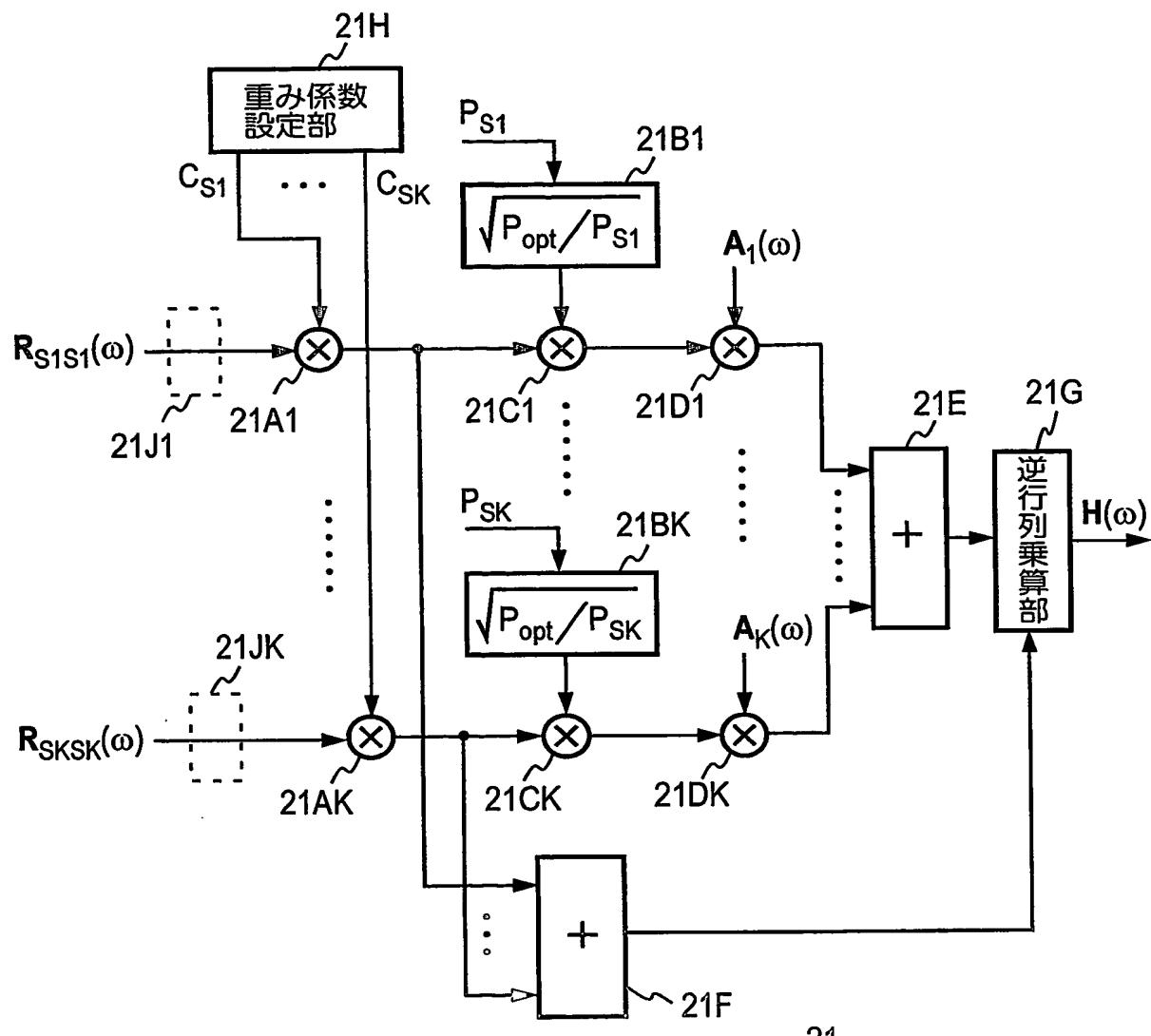
21

図4

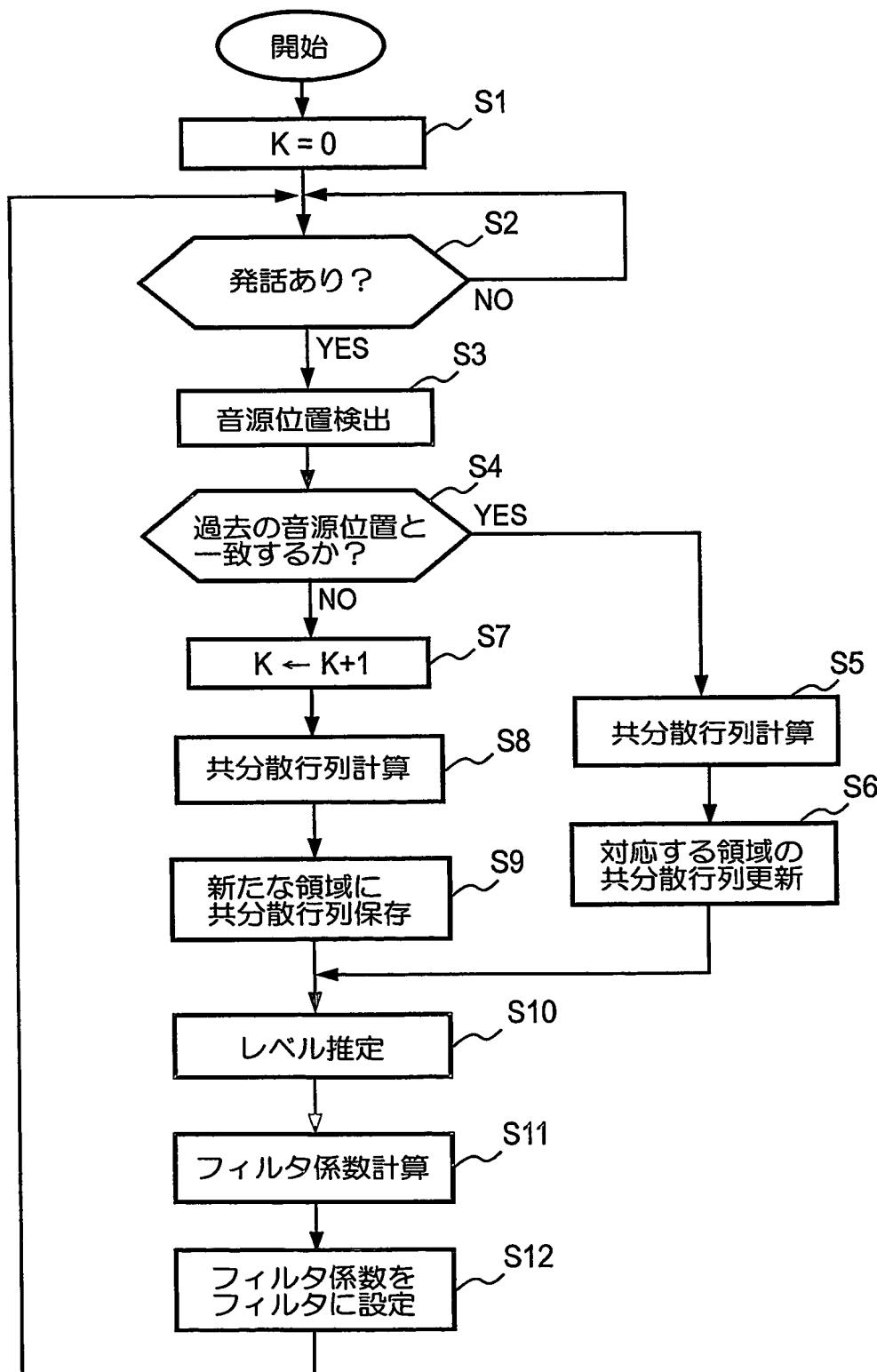


図5

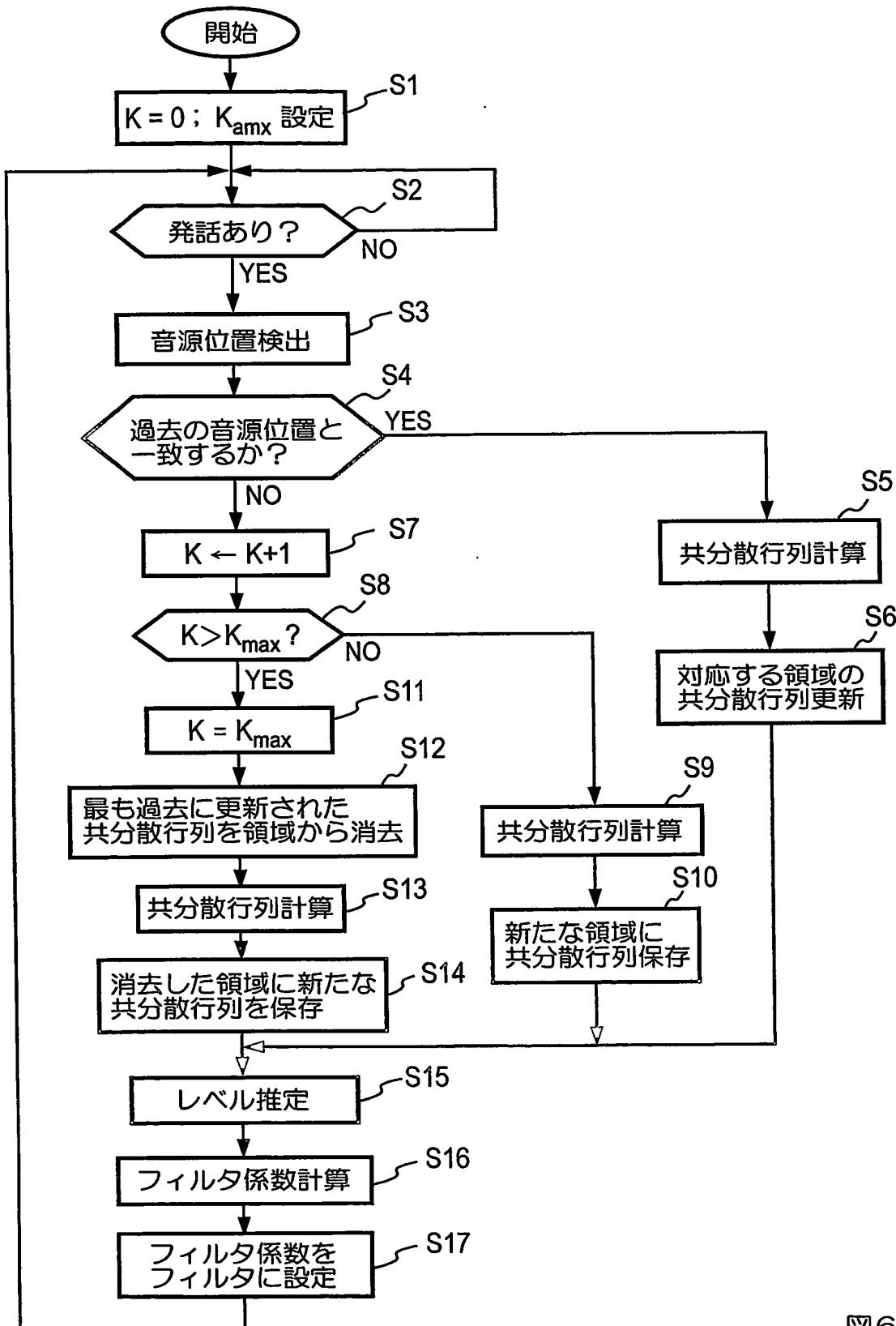


図6

7/18

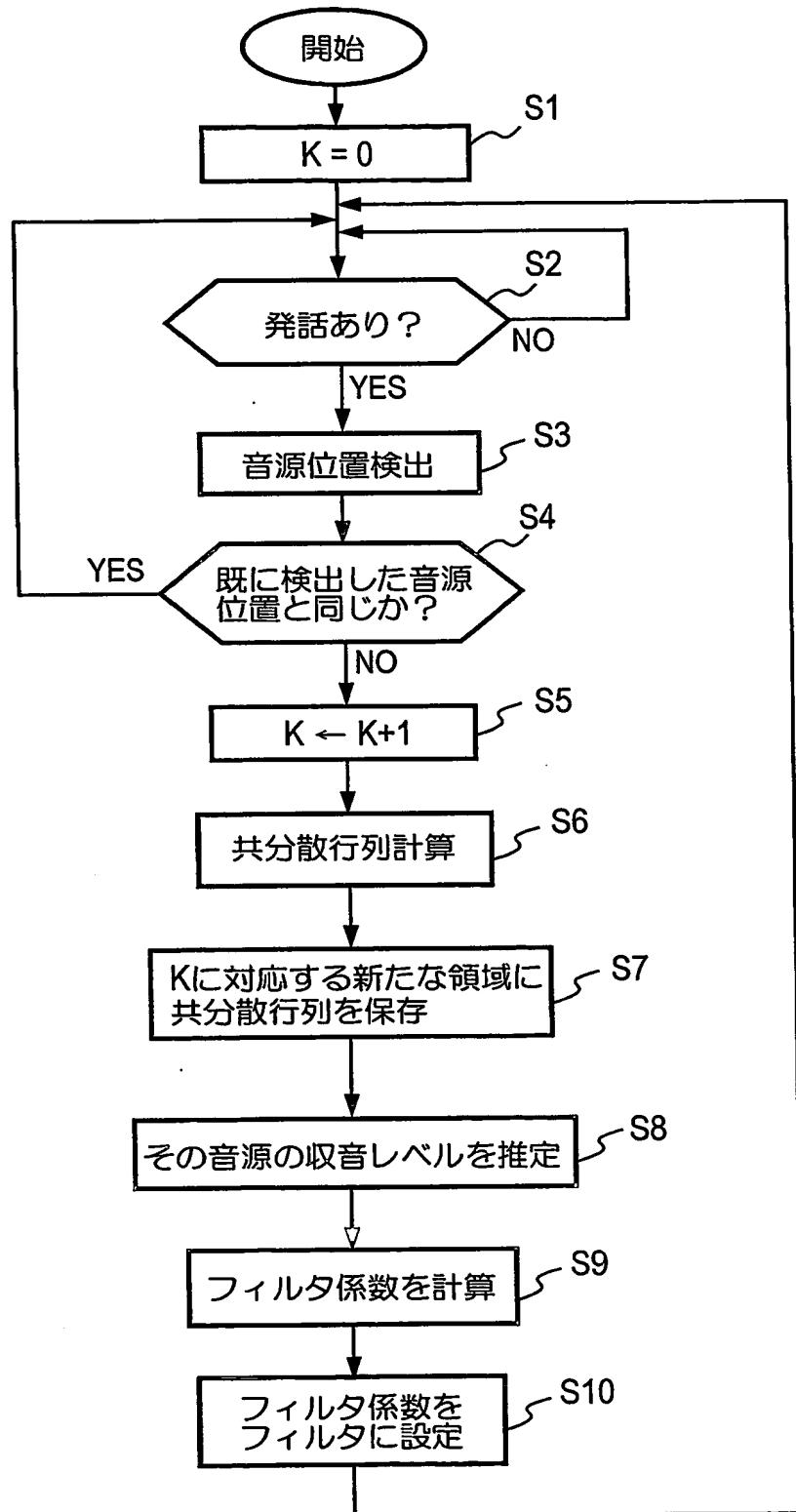


図7

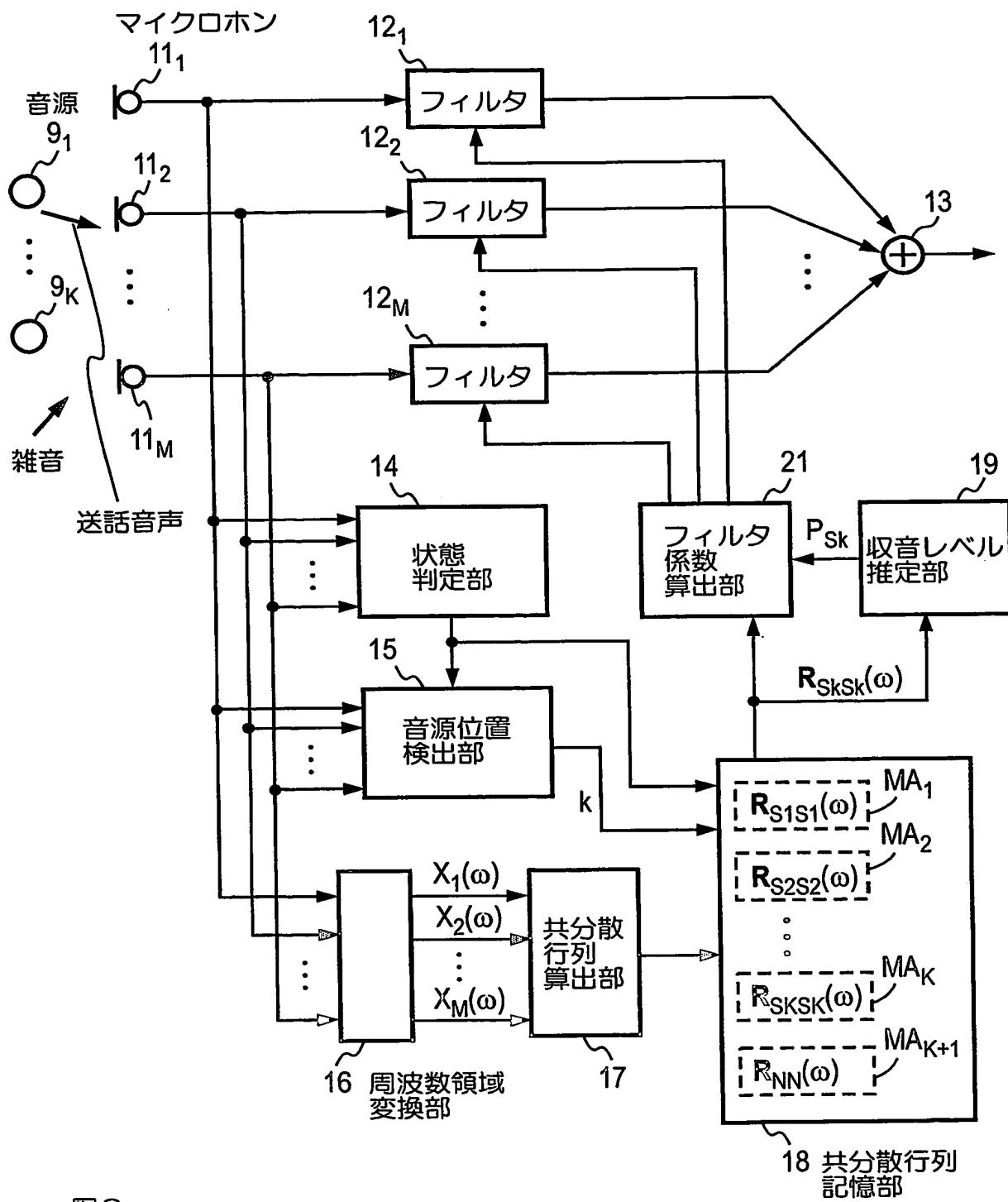


図8

9/18

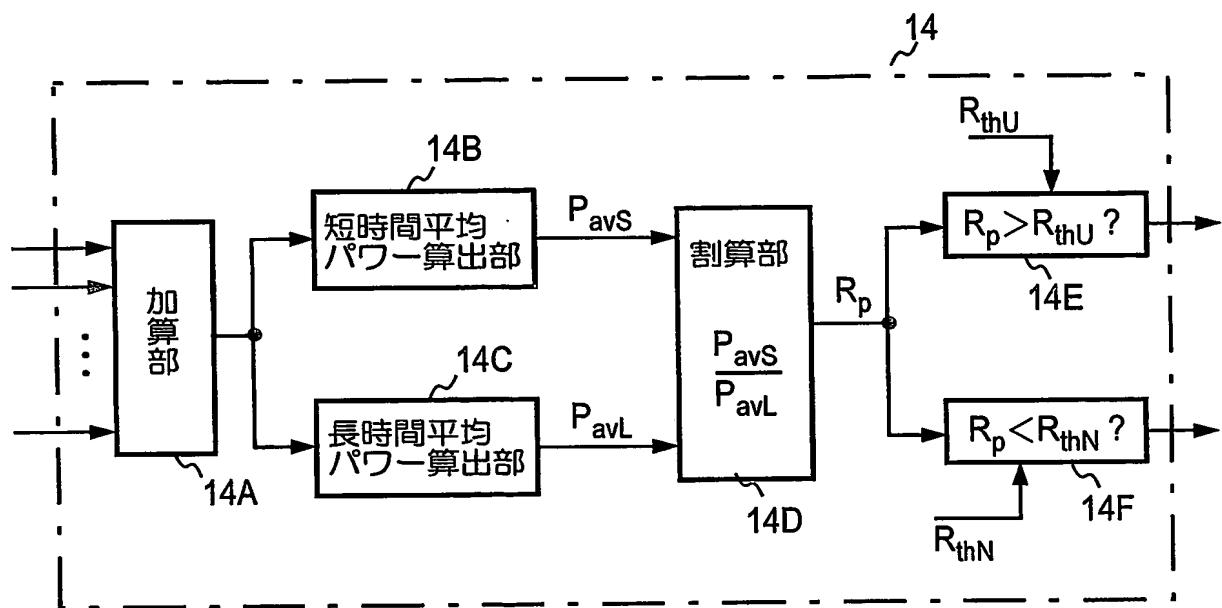
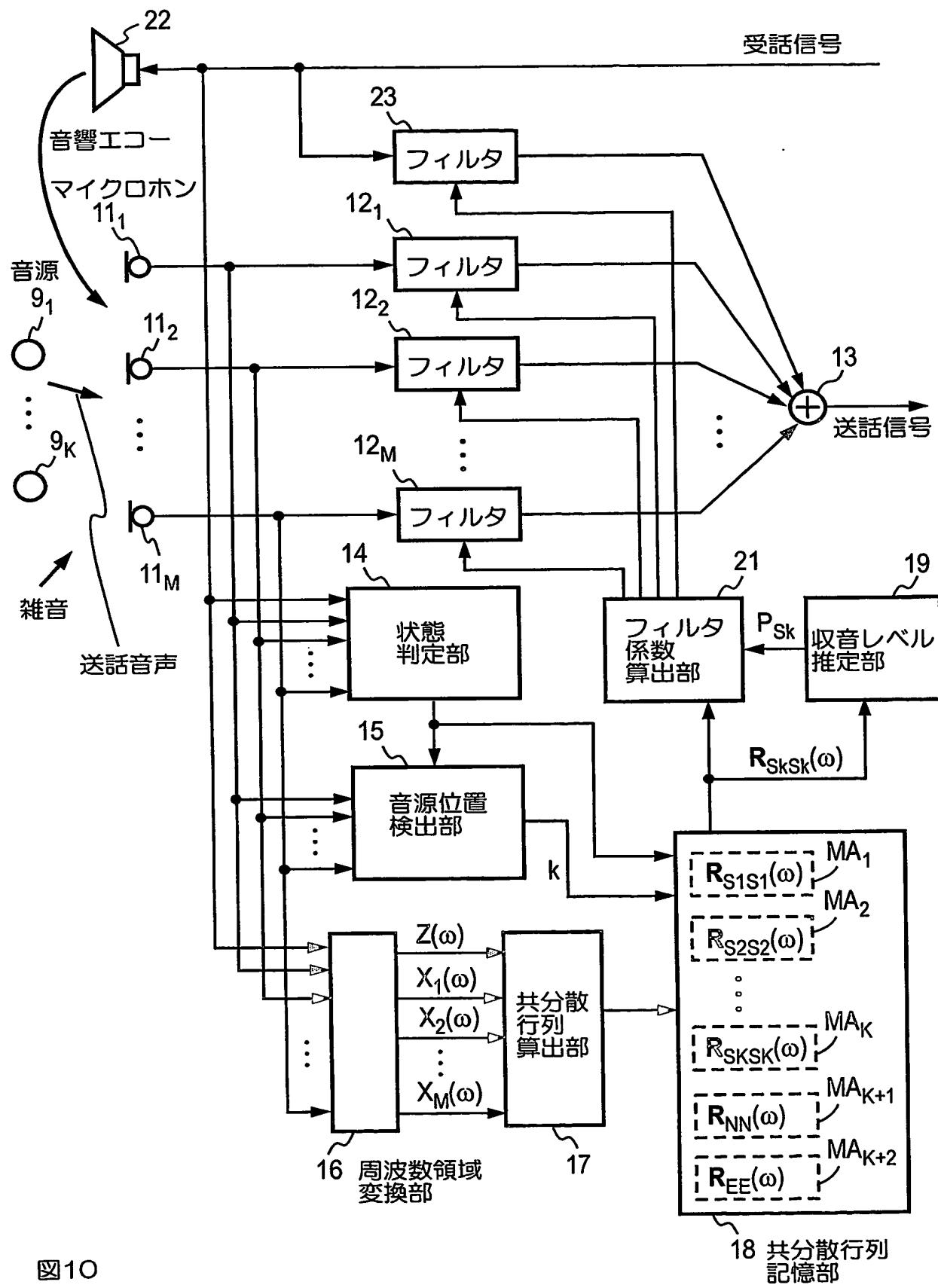


図9

10/18



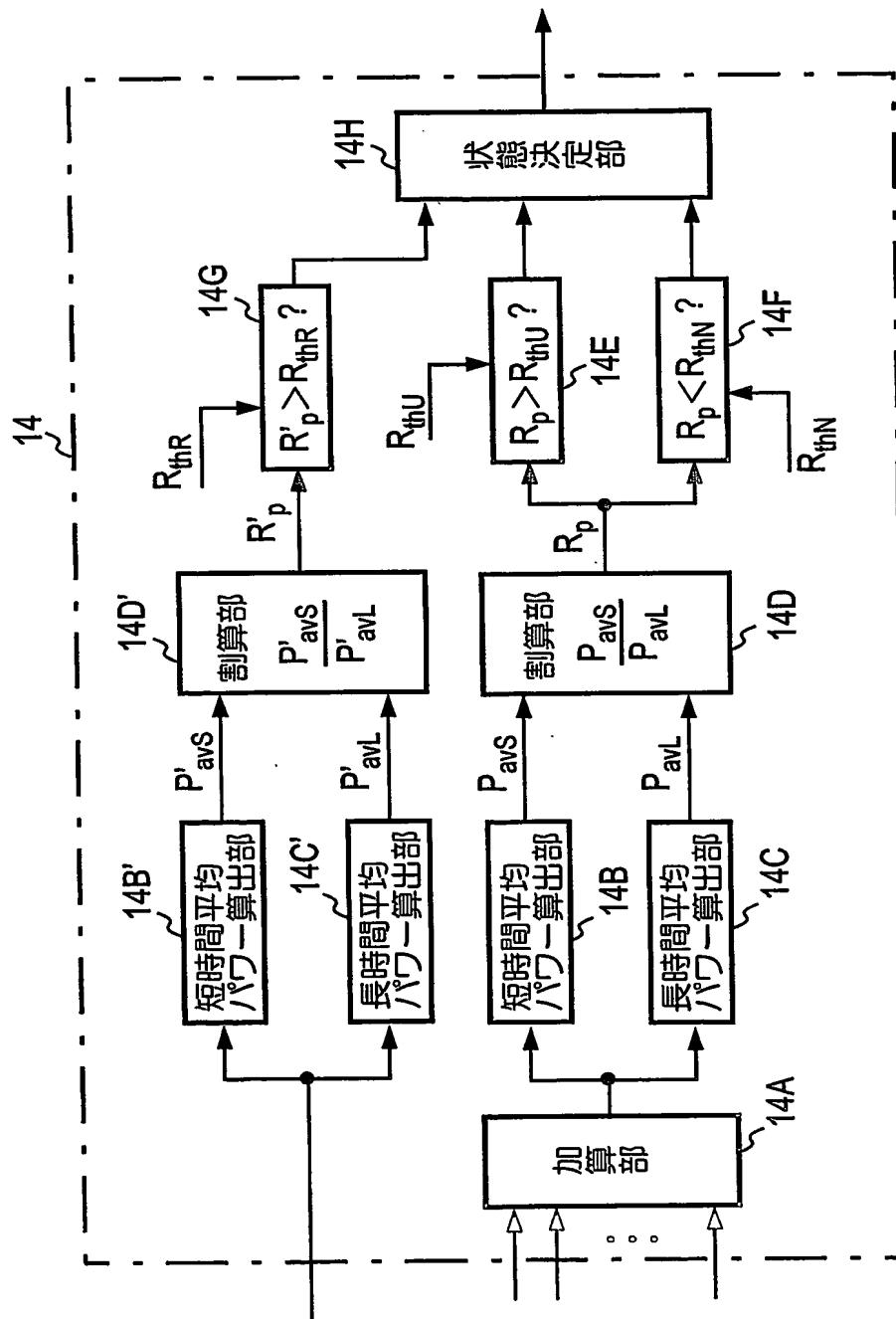


図11

12/18

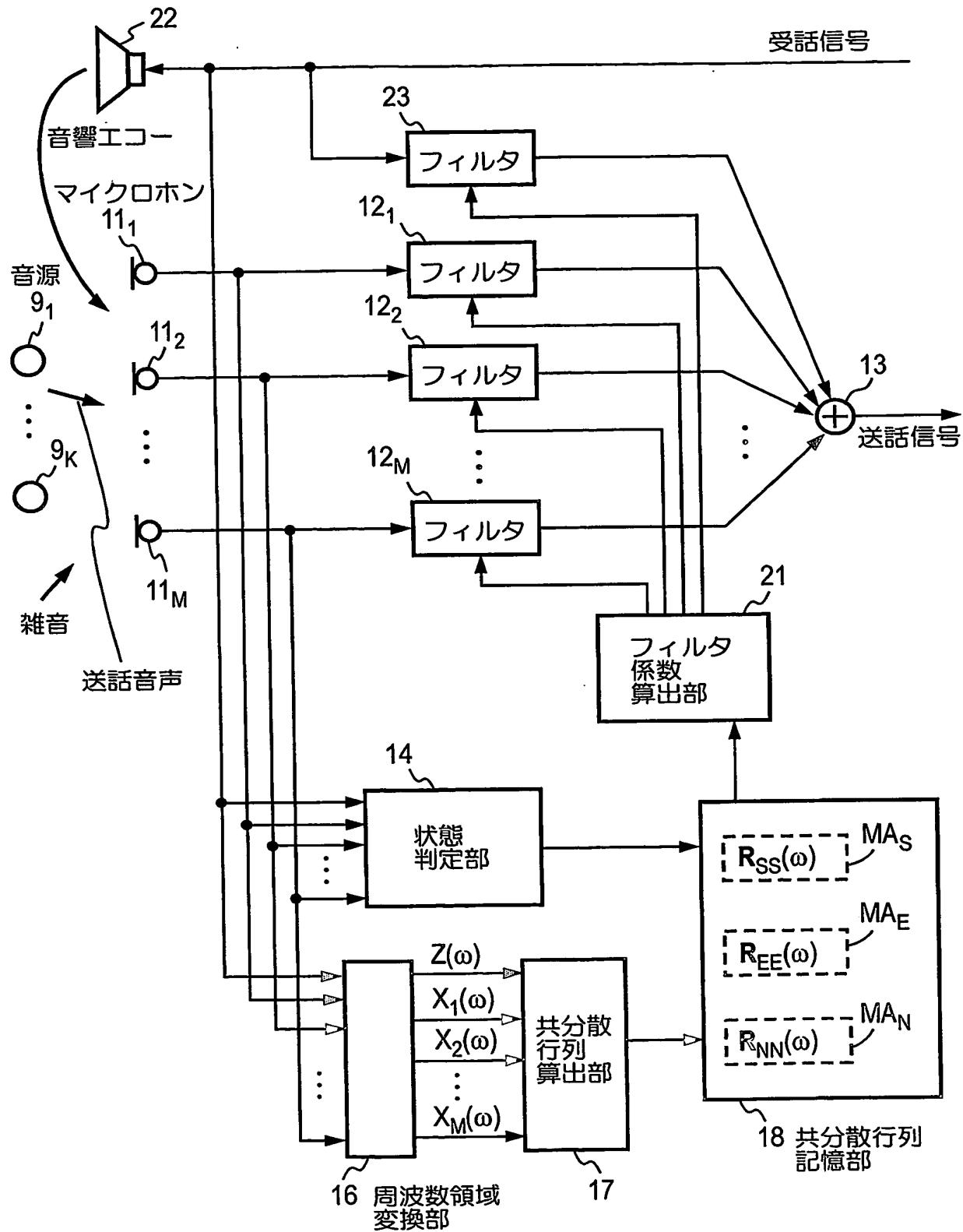


図12

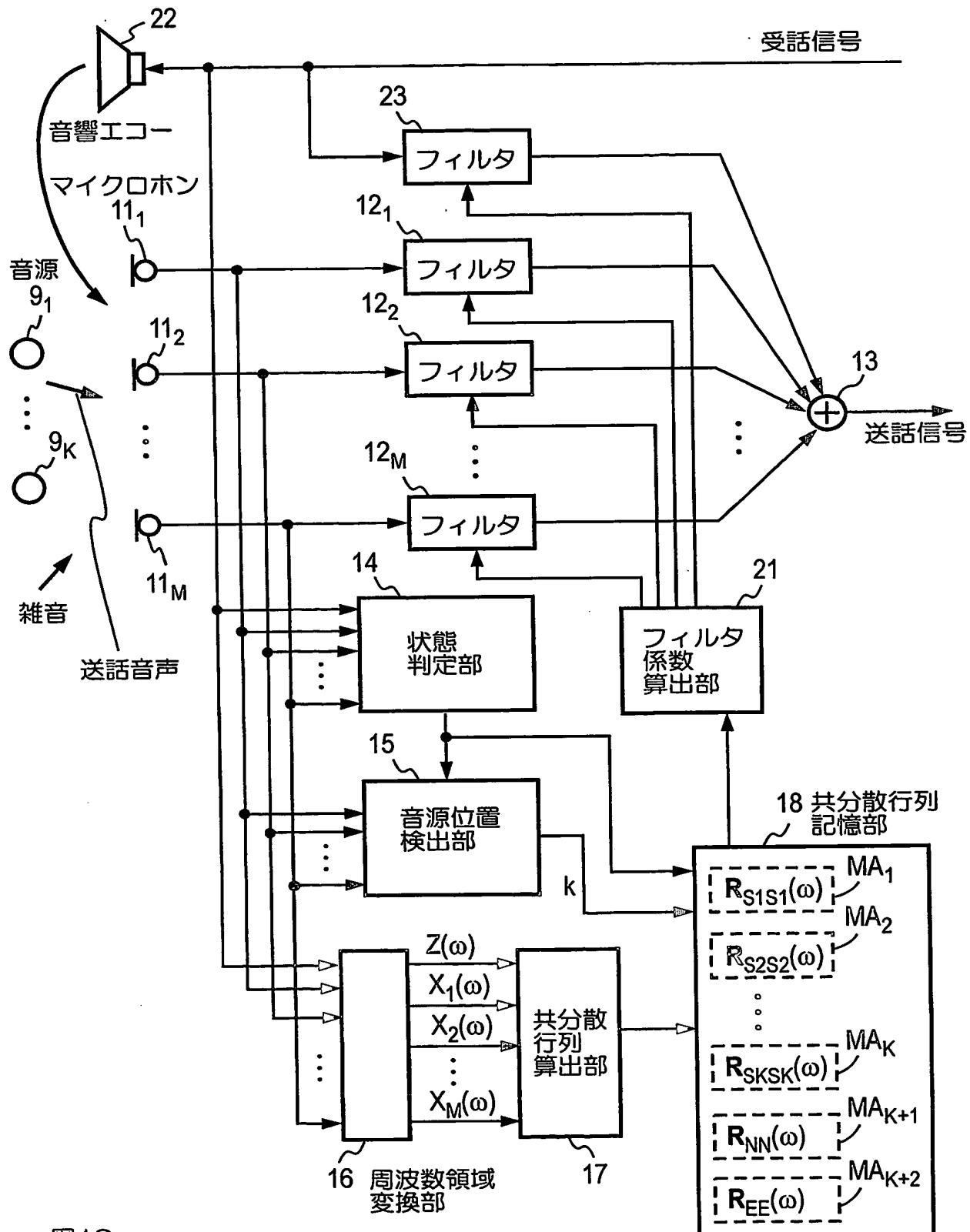


図13

14/18

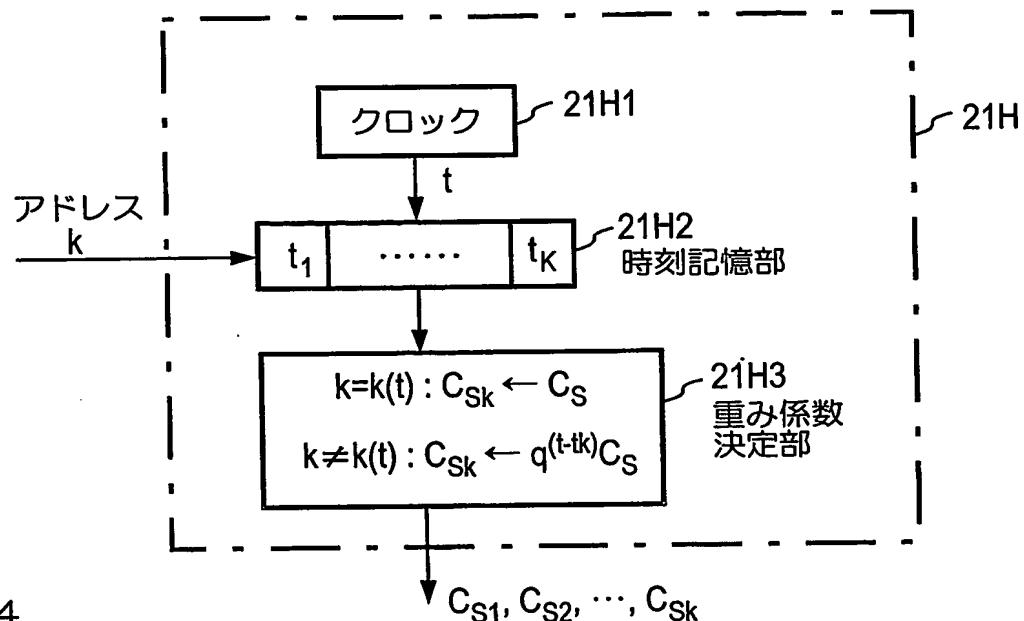


図14

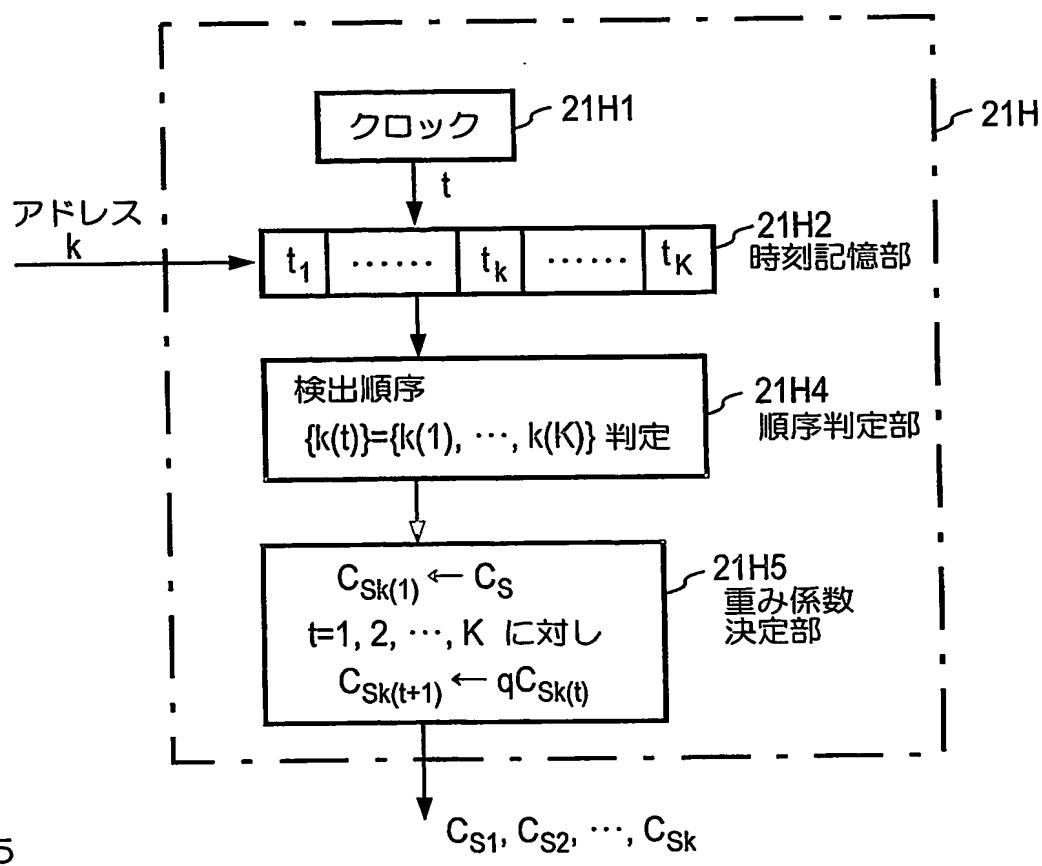


図15

15/18

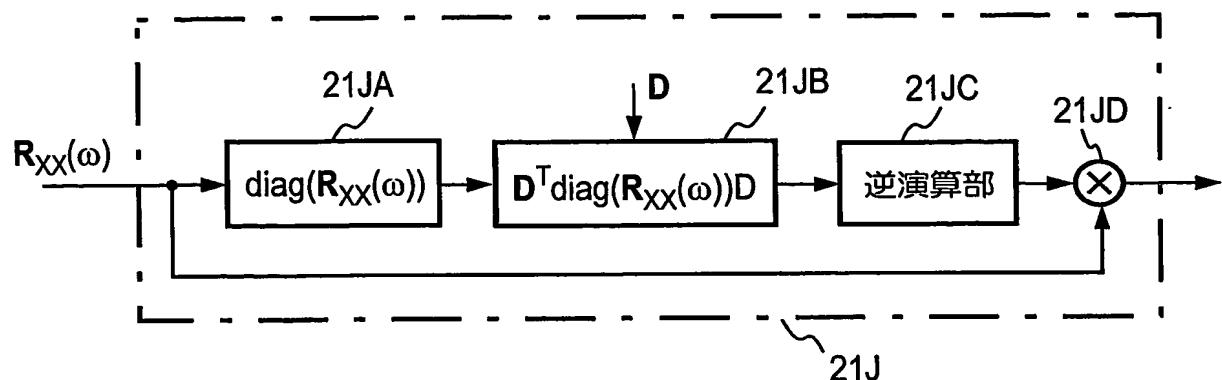


図16

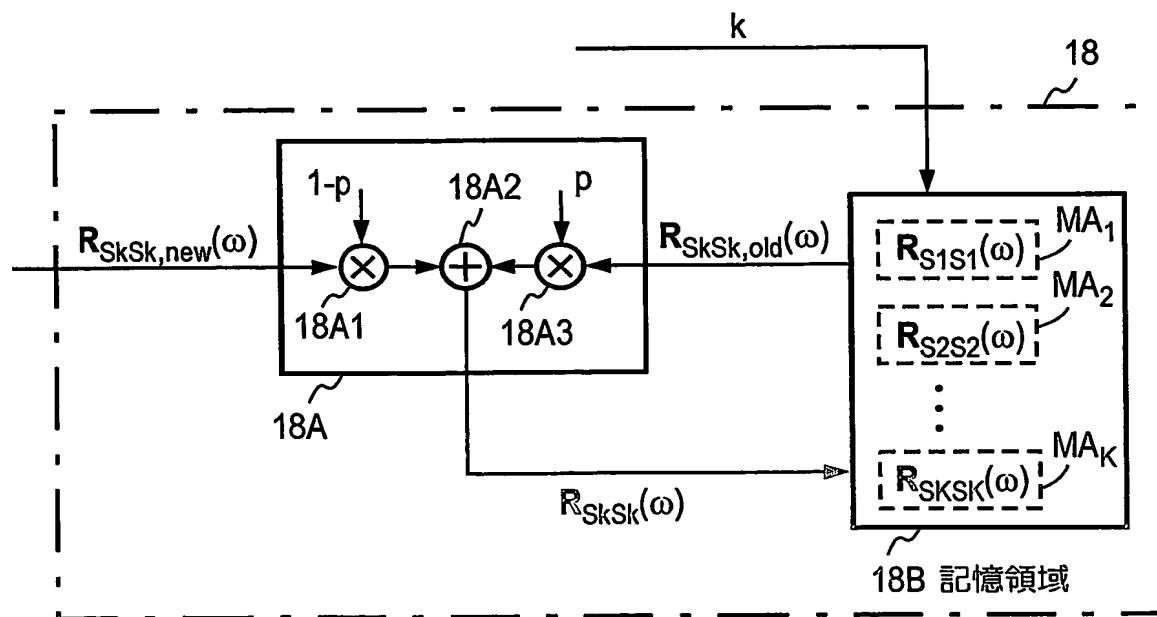


図17

16/18

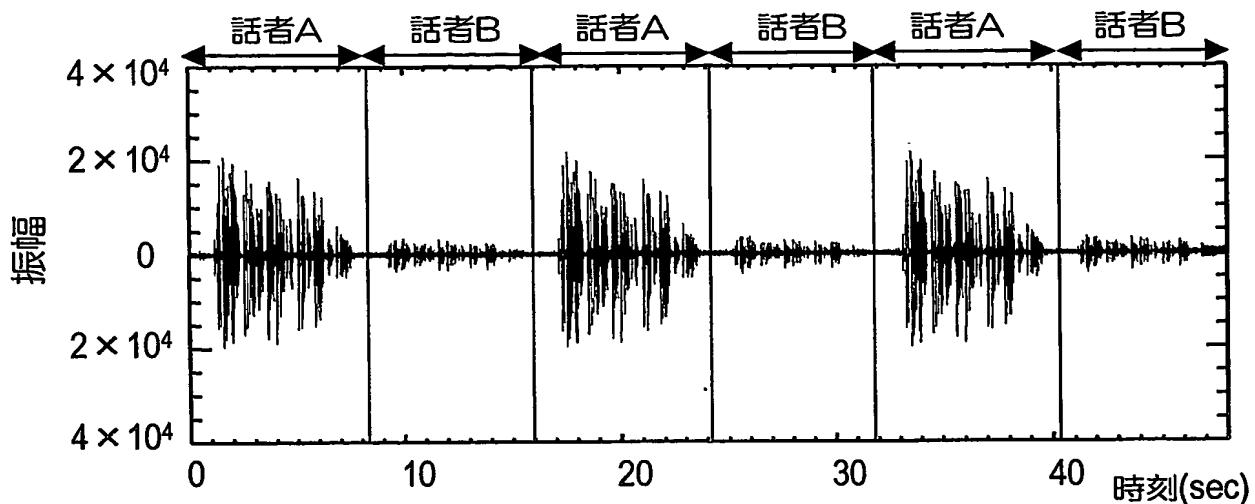


図18A 処理前

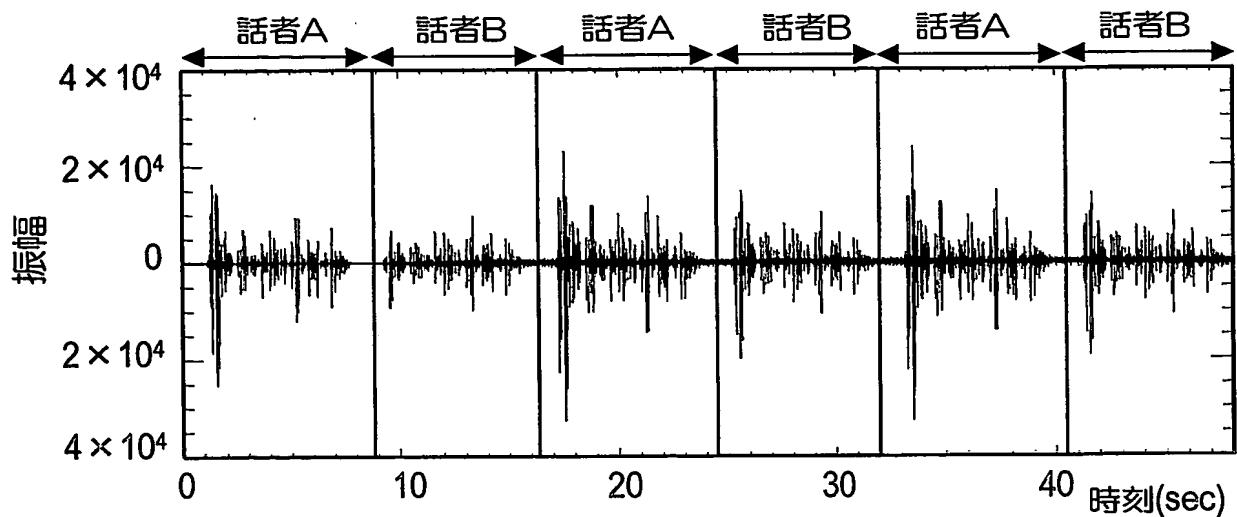
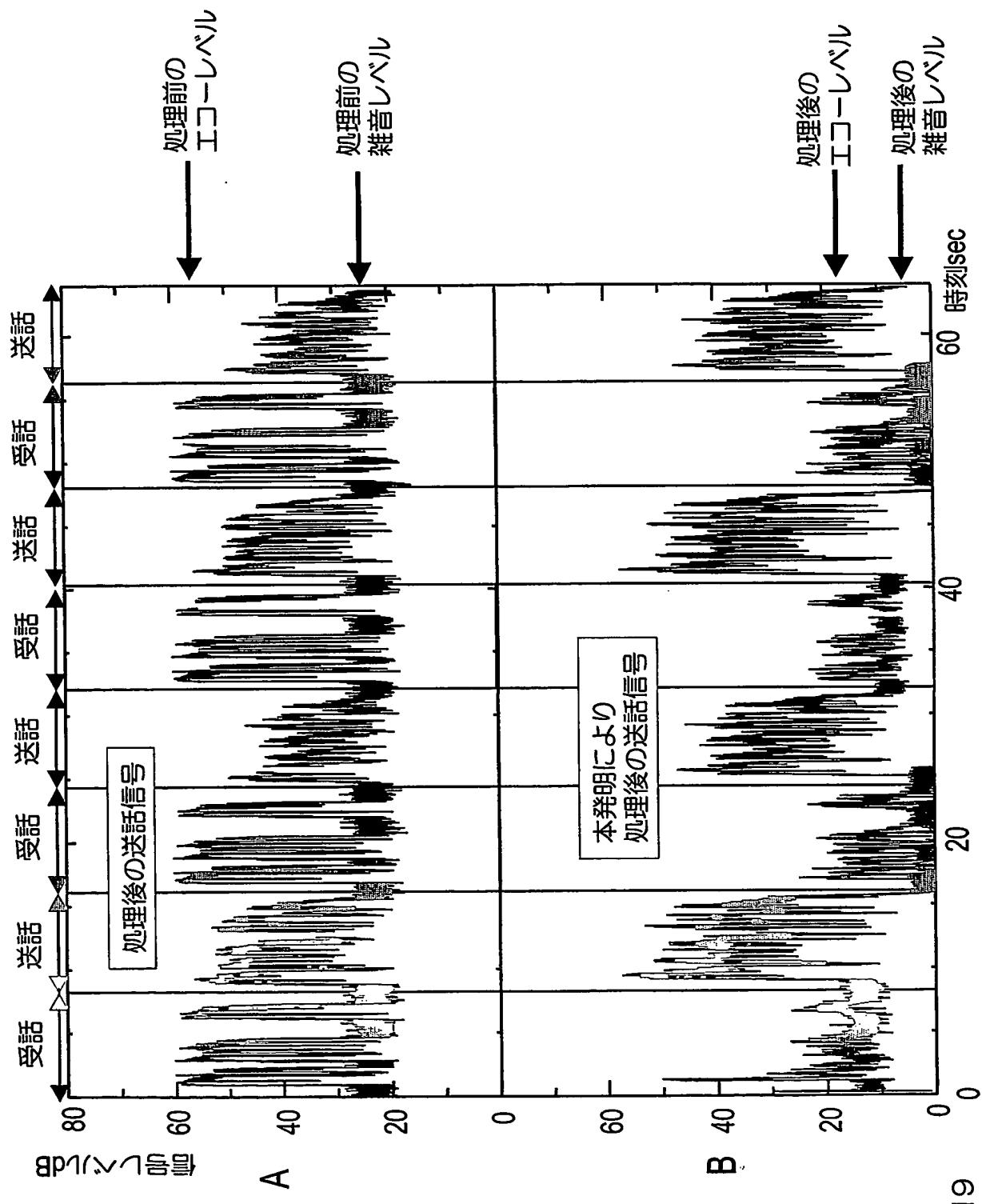


図18B 処理後



19

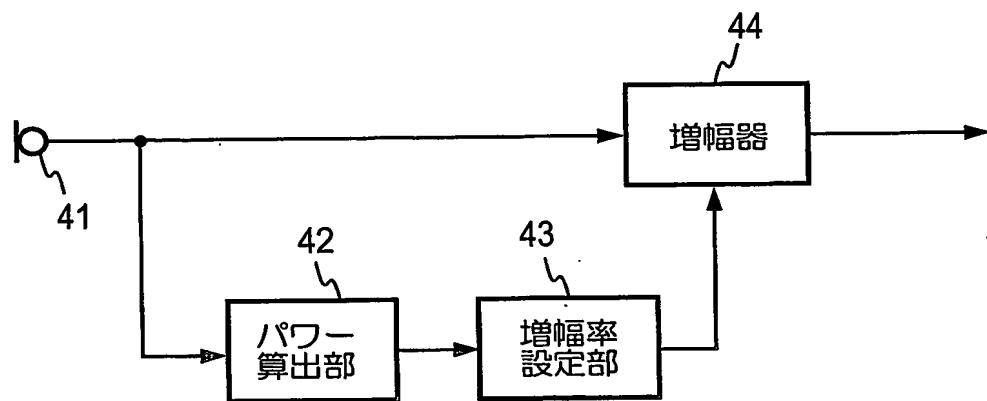


図20

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/001261

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ H04R3/00, H04B3/20, H03H21/00, H03H17/02, H04M1/60

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ H04R3/00, H04B3/20, H03H21/00, H03H17/02, H04M1/60

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1994-2004
Kokai Jitsuyo Shinan Koho 1971-2004 Jitsuyo Shinan Toroku Koho 1996-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 11-18194 A (Fujitsu Ltd.), 22 January, 1999 (22.01.99), Full text; all drawings & US 2003/21389 A1	1-21
A	JP 2002-335194 A (Yamaha Corp.), 22 November, 2002 (22.11.02), Full text; all drawings & US 6317501 B1 & US 2002/41693 A1	1-21
A	JP 10-190848 A (Lucent Technologies Inc.), 21 July, 1998 (21.07.98), Full text; all drawings & CA 2215592 A & US 5828756 A1 & EP 841799 A3	1-21

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	
"A"	document defining the general state of the art which is not considered to be of particular relevance
"E"	earlier application or patent but published on or after the international filing date
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O"	document referring to an oral disclosure, use, exhibition or other means
"P"	document published prior to the international filing date but later than the priority date claimed
"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&"	document member of the same patent family

Date of the actual completion of the international search
09 April, 2004 (09.04.04)Date of mailing of the international search report
18 May, 2004 (18.05.04)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/001261

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2002-149190 A (International Business Machines Corp.), 24 May, 2002 (24.05.02), Full text; all drawings & US 2002/136328 A1	1-21

A. 発明の属する分野の分類（国際特許分類（IPC））

Int. Cl' H04R3/00, H04B3/20, H03H21/00, H03H17/02, H04M1/60

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int. Cl' H04R3/00, H04B3/20, H03H21/00, H03H17/02, H04M1/60

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年

日本国公開実用新案公報 1971-2004年

日本国登録実用新案公報 1994-2004年

日本国実用新案登録公報 1996-2004年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 11-18194 A(富士通株式会社) 1999.01.22, 全文, 全図 & US 2003/21389 A1	1-21
A	JP 2002-335194 A(ヤマハ株式会社) 2002.11.22, 全文, 全図 & US 6317501 B1 & US 2002/41693 A1	1-21
A	JP 10-190848 A(レーセント テクノロジーズ インコーポレイテッド) 1998.07.21, 全文, 全図 & CA 2215592 A & US 5828756 A1 & EP 841799 A3	1-21
A	JP 2002-149190 A(インターナショナル・ビジネス・マシーンズ・コーポレーション) 2002.05.24, 全文, 全図 & US 2002/136328 A1	1-21

 C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

- 「A」特に関連のある文献ではなく、一般的技術水準を示すもの
- 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
- 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
- 「O」口頭による開示、使用、展示等に言及する文献
- 「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

- 「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
- 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
- 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
- 「&」同一パテントファミリー文献

国際調査を完了した日 09.04.2004	国際調査報告の発送日 18.5.2004
国際調査機関の名称及びあて先 日本国特許庁（ISA/JP） 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官（権限のある職員） 松澤 福三郎 電話番号 03-3581-1101 内線 3540